中国大学先修课程

线性代数

Linear Algebra

冯荣权　王殿军　杨晶　周俊

$$\begin{pmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{pmatrix}$$

高等教育出版社

VISIT…

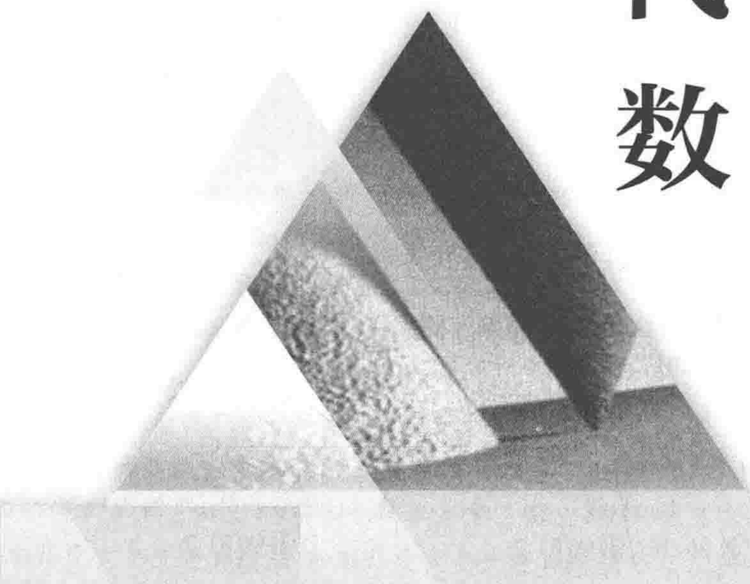
LANZAROTE Caliente.com

中国大学先修课程

Linear Algebra

线性代数

冯荣权　王殿军　杨晶　周俊



安徽师范大学图书馆
藏书

高等教育出版社·北京

内容提要

本书作为中国大学先修课程的教材，旨在使学生通过学习，理解线性代数中的基本概念、掌握线性代数中的基本理论和方法、接受以矩阵为代表的基本运算技能的训练，会处理线性代数中的常见问题，使学生得到比较系统的数学训练。全书共分八章，包括线性方程组、行列式、矩阵代数、向量空间、矩阵的特征值问题与二次型理论等。每一章内容后面都配备了形式多样的习题，书后附有绝大多数习题的答案和提示。本教材适合具有扎实的数学基础、学有余力，并希望提前选修大学数学基础课程的优秀高中生。

图书在版编目(CIP)数据

线性代数 / 冯荣权等编.-- 北京 ：高等教育出版社，2017.10
中国大学先修课程
ISBN 978-7-04-048260-7

Ⅰ.①线… Ⅱ.①冯… Ⅲ.①线性代数-高等学校-教材 Ⅳ.①O151.2

中国版本图书馆 CIP 数据核字(2017)第 185137 号

XIANXING DAISHU

策划编辑 李 茜　责任编辑 李 茜　封面设计 张申申　版式设计 马 云
插图绘制 黄云燕　责任校对 王 雨　责任印制 尤 静

出版发行 高等教育出版社
社　　址 北京市西城区德外大街 4 号
邮政编码 100120
印　　刷 北京明月印务有限责任公司
开　　本 889mm × 1194mm 1/16
印　　张 16.75
字　　数 410 千字
购书热线 010-58581118
咨询电话 400-810-0598
网　　址 http://www.hep.edu.cn
　　　　 http://www.hep.com.cn
网上订购 http://www.hepmall.com.cn
　　　　 http://www.hepmall.com
　　　　 http://www.hepmall.cn
版　　次 2017 年10月第 1 版
印　　次 2017 年10月第 1 次印刷
定　　价 48.00 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换
版权所有 侵权必究
物 料 号 48260-00

出版说明

《国家中长期教育改革和发展规划纲要(2010—2020年)》关于普通高中教育改革提出"深入推进课程改革。创造条件开设丰富多彩的选修课,为学生提供更多选择,促进学生全面而有个性的发展。"以及"推进培养模式多样化,满足不同潜质学生发展需要,探索发现和培养创新人才的途径"。关于人才培养体制改革提出"树立多样化人才观念,尊重个人选择,鼓励个性发展,不拘一格培养人才。树立系统培养观念,推进小学、中学、大学有机衔接"。

党的十八届三中全会对"深化教育领域综合改革"作出全面部署,对高中教育教学改革提出了新的更高要求。国务院于2014年发布了《国务院关于深化考试招生制度改革的实施意见》,明确要求启动高考综合改革试点。考试招生制度改革为教育综合改革带来了活力,使得学有余力的中学生能把更多的精力用于发展个人兴趣和特长,为在高中开设大学先修课程提供了必要条件。

为贯彻落实有关精神和要求,2014年3月,中国教育学会联合高等教育出版社共同发起并组织实施了"中国大学先修课程(CAP)试点项目"。本项目旨在探索加强高中与大学教育的衔接,促进拔尖创新人才培养的有效形式,为不断深化我国高中教育教学改革先行先试。项目邀请国内多所知名高校和教育科研机构的专家、学者,在学习、借鉴先进国际课程理念和经验的基础上,着力研制、开发一套适合我国高中教育实际、具有鲜明中国特色的大学先修课程,使学有余力的高中生能根据自身的兴趣和能力自主选择、自愿学习,提前接受大学的思维方式、学习方法,发展在学科专业学习和研究方面的潜能,帮助其为大学学习乃至未来的职业生涯做好准备。

本试点项目根据大学开课的特点,先期研发了8门课程,自2014年9月起在65所高中试点,到2016年第四次开课,试点学校增加到103所,选课总人数超过万人。同时,遵循能力导向的原则设计测评体系,给学生充分发挥的空间,鼓励学生通过推理、探究充分展示解题思路,并根据先进的教育理念和学科人才评测技术,借鉴国际大型测试项目的评分规则对结果进行评价。

试点期间,为配合试点项目课程研发,由中国教育学会委托国内知名大学和中学的专家为本系列课程编写了讲义。经过四个学期的试用,得到广大试点学校的积极反馈,对讲义内容适用性的提高起到关键作用,我们也希望这些经过实践锤炼的讲义形成的教材,能使更多的学生受益。本系列教材旨在将大学的学科教学内容比较完整地呈现给高中学生,在编写过程中着重考虑到以下两个方面:

1. 大学先修课程不是大学内容的简单照搬，教材的编写既考虑了大学的知识水准又考虑到了中学生的学习特点，引导学生尝试使用大学的自主学习方式来学习，鼓励培养主动学习的意识。

2. 大学先修课程的学习应建立在兴趣的基础上，由学生自由选择，目前除了数学、物理、文学、外语，我们的课程还拓展到中学课标没有涉及的经济学领域，未来还会不断有新的课程加入进来。

本项目尚处在试点阶段，在课程设置、教学内容、教学方法和测评方面虽然已经做出了很多尝试，并且取得了宝贵的经验，但在诸多方面的探索还在进行之中，距离完善还有相当的路程。希望广大的试点中学、教师、学生和家长能对本项目给予正确的理解与积极的支持，我们共同携手为高中教育的改进不懈努力。

为了体现教育公平，我们还在爱课程网上线了一批大学先修课程慕课(MOOC)，为没有条件在中学开设大学先修课的学校提供免费的教学和学习的平台，欢迎广大中学生来体验、学习。

中国大学先修课试点项目管理委员会办公室

二〇一六年五月

前　言

《线性代数》作为中国大学先修课程的教材之一，旨在使学生通过学习，理解线性代数中的基本概念、掌握线性代数中的基本理论和方法、接受以矩阵为代表的基本运算技能的训练，会处理线性代数中的常见问题，使学生得到比较系统的数学训练，进一步激发学生的学习兴趣、提升学生的科学素养，为进入下一阶段的学习做好知识储备和能力准备，为今后学习各类后继课程和进一步扩大数学知识面奠定必要的数学基础。

中国的大学先修课程教育刚刚起步，其教育教学模式还在探讨阶段。我们理解大学先修课程首先是大学课程，应该把它定位为既有大学学术标准和大学基本学业水平的，并且又符合中学生的数学基础和认知能力的衔接课程。所以这本教材在内容方面，选取了大学线性代数课程中的最基本的内容，主要包括：几何向量、空间的直线和平面、线性方程组的高斯（Gauss）消去法、行列式、矩阵运算、n 维向量空间、向量的线性相关性、内积与欧氏空间、矩阵的特征值与特征向量、二次型、几何变换等。因此本书适合具有扎实的数学基础、学有余力并希望提前先修大学基础课程的优秀中学生。

在内容处理上，需要特别说明以下几点。

（1）代数与几何相结合是本书的一个特点。代数学主要讨论代数运算的运算性质，而这些运算则是从具体的事例中抽象出来的，希望学习本课程的学生可以从学习中慢慢体会这一点。几何可以为许多代数概念提供原始模型，而代数可以为解决几何问题提供方法。所以在本书第一章我们给出了中学生所熟悉的空间几何模型，在最后一章，对之前讨论的代数方面的概念和理论给出了几何方面的对应和解释；借助矩阵变换给出了线性变换一个朴素的几何模型；并给出用代数方法解决空间中二次曲面分类的应用。同时，本书还给出了线性代数在其他领域应用的例子。本课程不仅在很多学科有重要应用，而且是后继课程的基础，同时它还担负着培养学生的综合运用所学的数学原理和技能分析问题和解决问题的能力、以抽象思维和逻辑推理为特征的理性思维能力和自主学习的能力，并可激发学生的创新精神。

（2）为了知识的完整，我们给出了几乎所有结论的证明，其中的部分证明对中学生来说难度有些大，任课老师在教学过程中可以灵活掌握。同时任课老师在课程的教学过程中，要突出数学思想方法的教学，加强数学应用能力的培养，淡化运算技巧的训练。另外学会形式化的表达是一项基本要求，但是不能只限于形式化的表达，要强调对数学本质的认识，不能将生动活泼的数学思维活动淹没在形式化的海洋里。因此，本书力图揭示线性代数中概念、理论、方法的发

展过程和本质。任课教师在教学中不仅要讲逻辑推理,还要讲道理,通过典型例子的分析和学生自主探索活动,使学生理解这些概念、理论逐步形成的过程,体会蕴涵在其中的思想方法。

(3) 本书中少量内容与部分高中课本内容相似,例如:第零章的前两节,第 8.1.1 节矩阵变换的定义和基本性质,§8.3 最小二乘问题,第 8.4.1 节平面中的二次曲线等,之所以保留这些内容,是出于知识结构的完整性考虑,也可使本书与高中教材衔接更加顺畅。在具体的授课过程中,教师可根据实际情况略讲或跳过这些部分。

(4) 学时建议:建议分两学期讲授,每学期各 36 学时(含习题讨论课),第一学期可讲授第零章至第四章,重点是线性方程组、行列式与矩阵的基本运算和求解,侧重具体计算。第二学期讲授第五章至第八章,主要内容是引入向量组的线性相关性理论,把之前的内容串接起来,并用其讨论更深的问题,如:线性方程组解的结构、矩阵的特征值问题、二次型问题等,侧重于理论分析与综合计算。如要一个学期讲授完,建议不少于 64 学时(不含习题讨论课)。

大学先修课程在国外已走过了几十年的历程,也有比较成功的经验。但在我国,大学先修课程的兴起却是近几年的事情。特别是线性代数课程,在美国 AP 课程中也是没有的。如何建设适合我国国情的大学先修课程,编写符合课程定位、能够达到课程目的的教材,一切都还在探索中。

在前期调研的基础上,本书参编人员在 2014 年 3 月至 7 月,经过多次研讨,确定了课程的定位及教材应包含的内容。教材中第一章至第六章的初稿于 2014 年 7 月底形成,并听取了 2014 年 8 月初在清华附中举办的中国大学先修课程试点项目首批实验学校教师培训的全体学员的意见和建议。根据反馈意见,编者对全书做了修改和校订,2015 年 1 月在清华附中又举办了中国大学先修课程试点项目实验学校第二期教师培训,期间就第一章至第六章教材内容及教学中的使用情况,征询了老师们的意见和建议。迄今为止,这部分内容已经经过几十所中学的多轮试用,并根据使用情况进行了修订。

本书的第零章、第七章和第八章的初稿在 2015 年 6 月底形成,并听取了中国大学先修课程试点项目实验学校第三期、第四期教师培训的全体学员的意见和建议。之后,编者根据反馈意见对全书做了修改、订正和重排。2016 年 3 月和 8 月,在清华附中举行了第一次和第二次 MOOCAP 教师工作坊研讨会,再一次征询了一线教师对教材的意见。迄今为止,该部分内容也已经过试用,并根据使用情况进行了修订。

编写一本合适的大学先修课程教材,对我们所有参编者都是一种挑战,尽管我们很努力,但由于学识、认识及经验所限,书中难免会有不足,希望使用本教材的老师和学生提出宝贵的意见和建议,您的关注是我们做好此项工作的最大动力。

编 者

2016 年 11 月于北京

目　录

第零章　预备知识 …… 1

§0.1　2、3 阶行列式 …… 1
§0.2　行列式的性质 …… 4
§0.3　求和号与连乘符号 …… 9
§0.4　数域 …… 12
习题零 …… 13

第一章　空间与向量 …… 14

§1.1　向量与空间直角坐标系 …… 14
1.1.1　向量的基本概念 …… 14
1.1.2　向量的线性运算及投影 …… 15
1.1.3　空间直角坐标系 …… 17
§1.2　向量的内积、外积与混合积 …… 19
1.2.1　内积 …… 19
1.2.2　外积 …… 21
1.2.3　混合积 …… 23
1.2.4　向量间的关系 …… 24
§1.3　空间平面及其方程 …… 25
1.3.1　平面的点法式方程 …… 25
1.3.2　平面的一般式方程 …… 26
1.3.3　平面的截距式方程 …… 27
1.3.4　平面的三点式方程 …… 28
1.3.5　同轴平面束 …… 28
§1.4　空间直线及其方程 …… 29
1.4.1　直线的点向式方程与参数式方程 …… 29
1.4.2　直线的一般式方程 …… 30
§1.5　位置关系、夹角与距离 …… 32
1.5.1　两平面间的关系 …… 32
1.5.2　直线与平面间的关系 …… 32
1.5.3　两直线间的关系 …… 33
1.5.4　直线和平面相互间的夹角 …… 34
1.5.5　距离 …… 35
习题一 …… 37

第二章　线性方程组 …… 40

§2.1　线性方程组的解法 …… 40
§2.2　线性方程组解的情况的判定 …… 45
§2.3　齐次线性方程组 …… 48
习题二 …… 50

第三章　行列式 …… 52

§3.1　n 元排列 …… 52
§3.2　n 阶行列式 …… 54
§3.3　行列式的性质 …… 56
§3.4　行列式的展开公式 …… 61
§3.5　克拉默(Cramer)法则 …… 68

习题三 …… 70

第四章 矩阵代数 …… 72

§4.1 矩阵及其运算 …… 72
4.1.1 加法和数乘 …… 72
4.1.2 乘法 …… 73
4.1.3 转置 …… 75
4.1.4 矩阵的运算与行列式 …… 76
§4.2 分块矩阵 …… 78
4.2.1 分块矩阵 …… 78
4.2.2 初等矩阵与初等变换 …… 81
§4.3 可逆矩阵 …… 83
习题四 …… 88

第五章 向量空间 …… 90

§5.1 n 维向量空间 …… 90
5.1.1 n 维向量空间与子空间 …… 91
5.1.2 向量组的线性组合与线性表出 …… 91
§5.2 向量组的线性相关性 …… 93
§5.3 向量组的秩 …… 97
5.3.1 两个向量组的关系 …… 97
5.3.2 极大线性无关组 …… 98
5.3.3 向量组的秩 …… 99
5.3.4 向量(子)空间的基与维数 …… 100
§5.4 矩阵的秩 …… 101
5.4.1 矩阵的三种秩 …… 101
5.4.2 矩阵秩的计算 …… 102
5.4.3 三种秩的统一 …… 103
5.4.4 矩阵秩的性质 …… 103
§5.5 线性方程组的解理论 …… 104
5.5.1 齐次线性方程组有非零解的条件 …… 104
5.5.2 齐次线性方程组的解空间与基础解系 …… 105
5.5.3 非齐次线性方程组有解的条件 …… 109
5.5.4 非齐次线性方程组有解时解的结构 …… 110
5.5.5 $\mathbb{R}^3$ 中三个平面的位置关系 …… 113
§5.6 欧氏空间 …… 114
5.6.1 内积与欧氏空间 …… 114
5.6.2 内积与度量 …… 115
5.6.3 一个矩阵秩的例子 …… 116
5.6.4 标准正交基和正交矩阵 …… 117
5.6.5 施密特(Schmidt)正交化方法 …… 118
5.6.6 可逆实矩阵的 QR 分解 …… 120
习题五 …… 122

第六章 矩阵的特征值问题 …… 124

§6.1 矩阵的特征值与特征向量 …… 124
6.1.1 基本概念 …… 124
6.1.2 特征值与特征向量的计算方法 …… 125
6.1.3 特征多项式及其性质 …… 127
6.1.4 特征向量的性质与特征子空间 …… 128
§6.2 矩阵的相似与对角化 …… 130
6.2.1 矩阵的相似 …… 130
6.2.2 矩阵的对角化问题 …… 131
6.2.3 矩阵对角化的应用与例 …… 134
6.2.4 特征值理论的几个应用 …… 135
§6.3 实对称阵的对角化 …… 141
6.3.1 实对称阵的特征值与特征向量的性质 …… 141
6.3.2 实对称阵的对角化 …… 143
6.3.3 实对称阵的对角化(主轴化)方法 …… 144
习题六 …… 146

第七章 二次型 …… 149

§7.1 二次型和它的标准形 …… 149
§7.2 复数域上二次型的规范形 …… 160
§7.3 实数域上二次型的规范形 …… 162

§7.4 正定二次型和正定矩阵 …… 165
习题七 …… 169

第八章 几何意义与应用专题 …… 172

§8.1 矩阵与变换 …… 172
8.1.1 矩阵映射与矩阵变换的定义 …… 172
8.1.2 矩阵映射的性质 …… 173
8.1.3 $\mathbb{R}^2$与$\mathbb{R}^3$中几类特殊的矩阵变换 …… 175
8.1.4 矩阵映射的复合与矩阵乘法 …… 180
8.1.5 变换的不变量与特征值理论 …… 181
8.1.6 坐标系替换与矩阵相似 …… 183
8.1.7 正交变换 …… 187
§8.2 行列式的几何意义 …… 191
8.2.1 二阶、三阶行列式的几何意义 … 191
8.2.2 矩阵变换与行列式 …… 193
8.2.3 一般行列式的几何意义 …… 195
§8.3 最小二乘问题 …… 197
8.3.1 近似解的标准——最小距离 …… 197
8.3.2 最小二乘解的应用 …… 200
*8.3.3 最小二乘问题的微积分推导 …… 203
§8.4 二次曲线与二次曲面 …… 204
8.4.1 $\mathbb{R}^2$中的二次曲线及其标准方程 …… 204
8.4.2 二次曲线一般方程的化简与分类 …… 207
8.4.3 3维空间中的二次曲面 …… 213
习题八 …… 220

部分习题参考答案 …… 224

致谢 …… 251

第零章 预备知识

§0.1 2、3阶行列式

考察下面二元一次方程组

$$\begin{cases}a_{11}x+a_{12}y=b_1,\\ a_{21}x+a_{22}y=b_2,\end{cases}\tag{0.1}$$

其中 x,y 为未知量，$a_{11},a_{12},a_{21},a_{22}$ 为未知量系数，b_1,b_2 为常数项.不妨设 a_{11},a_{21} 均不为 0，则第 1 个方程两边乘以 a_{21}，第 2 个方程两边乘以 a_{11} 得

$$\begin{cases}a_{11}a_{21}x+a_{12}a_{21}y=b_1a_{21},\\ a_{11}a_{21}x+a_{11}a_{22}y=b_2a_{11}.\end{cases}$$

第 2 个方程减去第 1 个方程得到

$$(a_{11}a_{22}-a_{12}a_{21})y=b_2a_{11}-b_1a_{21}.$$

若 $a_{11}a_{22}-a_{12}a_{21}\neq 0$，则有

$$y=\frac{b_2a_{11}-b_1a_{21}}{a_{11}a_{22}-a_{12}a_{21}}.$$

将上述 y 的值代入第 1 个方程，得到

$$a_{11}x=b_1-\frac{b_2a_{11}-b_1a_{21}}{a_{11}a_{22}-a_{12}a_{21}}\cdot a_{12}=\frac{a_{11}(b_1a_{22}-b_2a_{12})}{a_{11}a_{22}-a_{12}a_{21}},$$

故

$$x=\frac{b_1a_{22}-b_2a_{12}}{a_{11}a_{22}-a_{12}a_{21}}.$$

所以，若 $a_{11}a_{22}-a_{12}a_{21}\neq 0$，则上述二元一次方程组的解为

$$x=\frac{b_1a_{22}-b_2a_{12}}{a_{11}a_{22}-a_{12}a_{21}},\quad y=\frac{b_2a_{11}-b_1a_{21}}{a_{11}a_{22}-a_{12}a_{21}}.\tag{0.2}$$

类似地,对于下面的三元一次方程组

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1, \\ a_{21}x + a_{22}y + a_{23}z = b_2, \\ a_{31}x + a_{32}y + a_{33}z = b_3, \end{cases} \tag{0.3}$$

若 $a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}\neq 0$,则可得方程组(0.3)的解为

$$\begin{cases} x = \dfrac{b_1a_{22}a_{33}+b_2a_{13}a_{32}+b_3a_{12}a_{23}-b_1a_{23}a_{32}-b_2a_{12}a_{33}-b_3a_{13}a_{22}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}, \\ y = \dfrac{b_1a_{23}a_{31}+b_2a_{11}a_{33}+b_3a_{13}a_{21}-b_1a_{21}a_{33}-b_2a_{13}a_{31}-b_3a_{11}a_{23}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}, \\ z = \dfrac{b_1a_{21}a_{32}+b_2a_{12}a_{31}+b_3a_{11}a_{22}-b_1a_{22}a_{31}-b_2a_{11}a_{32}-b_3a_{12}a_{21}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}. \end{cases} \tag{0.4}$$

注意到上面方程组的解中分子、分母的形式,为了便于表述,我们引进记号如下:令

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

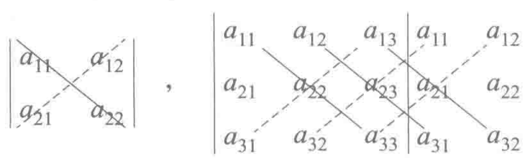
称其为一个 **2 阶行列式**.类似地,令

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31},$$

称其为一个 **3 阶行列式**.

行列式中每一横排称为它的行,每一竖排称为它的列.一个 2(或 3)阶行列式用一个 2(或 3)行 2(或 3)列的数表,左右两边再加上竖线来表示.数表中出现的数也称为此行列式的元素.

行列式中从左上角到右下角的线称为其**主对角线**.从右上角到左下角的线称为其**副对角线**.所以 2 阶行列式等于它的主对角线上两个元素的乘积减去副对角线上两个元素的乘积,而 3 阶行列式等于主对角线或平行于主对角线的 3 条线上元素的乘积之和再减去副对角线或平行于副对角线的 3 条线上元素乘积之和.具体表示如下



(0.5)

其中实线为主对角线(或与主对角线平行的线),虚线为副对角线(或与副对角线平行的线).

利用行列式表示,二元一次方程组(0.1)的解(0.2)可以写成

$$x=\frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}},\quad y=\frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \tag{0.6}$$

三元一次方程组(0.3)的解(0.4)可以写成

$$x=\frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}},\quad y=\frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}},\quad z=\frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}. \tag{0.7}$$

把行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

称为三元一次方程组(0.3)的系数行列式.所以当方程组(0.3)的系数行列式不为0时,它的解为

$$x=\frac{D_1}{D},\quad y=\frac{D_2}{D},\quad z=\frac{D_3}{D},$$

其中 D_i 是把系数行列式 D 的第 i 列换成常数项这一列得到的行列式.自然,这个结论对二元一次方程组(0.1)也成立.

例 0.1　解三元一次方程组

$$\begin{cases} x+3y+z=2, \\ 3x+4y+2z=9, \\ -x-5y+4z=10. \end{cases}$$

解　此方程组的系数行列式为

$$\begin{aligned} D=\begin{vmatrix} 1 & 3 & 1 \\ 3 & 4 & 2 \\ -1 & -5 & 4 \end{vmatrix} &= 1\times4\times4+3\times2\times(-1)+1\times3 \\ &\quad \times(-5)-1\times2\times(-5)-3\times3\times4-1\times4\times(-1) \\ &=-27\neq0. \end{aligned}$$

再用常数项列依次替换上面行列式中的列可得

$$D_1=\begin{vmatrix}2&3&1\\9&4&2\\10&-5&4\end{vmatrix}=-81,\quad D_2=\begin{vmatrix}1&2&1\\3&9&2\\-1&10&4\end{vmatrix}=27,$$

$$D_3=\begin{vmatrix}1&3&2\\3&4&9\\-1&-5&10\end{vmatrix}=-54.$$

所以 $x=\frac{D_1}{D}=3, y=\frac{D_2}{D}=-1, z=\frac{D_3}{D}=2$.即方程组的解为

$$\begin{cases}x=3,\\y=-1,\\z=2.\end{cases}$$

§0.2 行列式的性质

2 阶行列式是 2 项的代数和,每一项都是取自不同行不同列的 2 个元素的乘积.3 阶行列式是 6 项的代数和,每一项都是取自不同行不同列的 3 个元素的乘积.所以直接展开计算的运算量还是较大的.本节先给出行列式的性质.由性质出发可以简化行列式的计算.先考察 2 阶行列式.

性质 0.1 行列互换,行列式的值不变.即

$$\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}=\begin{vmatrix}a_{11}&a_{21}\\a_{12}&a_{22}\end{vmatrix}.$$

证明 由行列式定义,等式两边都等于 $a_{11}a_{22}-a_{12}a_{21}$. □

由性质 0.1 知行列式中行与列具有平等的性质.所以下面我们只对行来说明行列式的性质.自然行列式对列也具有相应的性质.

性质 0.2 两行互换,行列式的值变号.即

$$\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}=-\begin{vmatrix}a_{21}&a_{22}\\a_{11}&a_{12}\end{vmatrix}.$$

证明 $\begin{vmatrix}a_{21}&a_{22}\\a_{11}&a_{12}\end{vmatrix}=a_{21}a_{12}-a_{11}a_{22}$,所以上式两端都等于 $a_{11}a_{22}-a_{12}a_{21}$. □

由此性质可知,若行列式的性质对某一行成立,则对任意一行也成立.所以下面性质只对第一行来说明并证明.

性质 0.3 若行列式中某行每个元素分成两个元素之和,则该行列式可关于该行拆开成两个行列式之和.拆开时其他各行均保持不变.即

$$\begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

证明 $\begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21} & a_{22} \end{vmatrix} = (a_{11}+b_{11})a_{22}-(a_{12}+b_{12})a_{21}$

$$=(a_{11}a_{22}-a_{12}a_{21})+(b_{11}a_{22}-b_{12}a_{21})$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$ □

性质 0.4 行列式中某行有公因子 k 时,k 可以提出行列式外.即

$$\begin{vmatrix} ka_{11} & ka_{12} \\ a_{21} & a_{22} \end{vmatrix} = k\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

证明 $\begin{vmatrix} ka_{11} & ka_{12} \\ a_{21} & a_{22} \end{vmatrix} = (ka_{11})a_{22}-(ka_{12})a_{21} = k(a_{11}a_{22}-a_{12}a_{21}) = k\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$ □

性质 0.5 行列式中某一行加上另一行的 k 倍后,其值不变,即

$$\begin{vmatrix} a_{11}+ka_{21} & a_{12}+ka_{22} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

证明 $\begin{vmatrix} a_{11}+ka_{21} & a_{12}+ka_{22} \\ a_{21} & a_{22} \end{vmatrix} = (a_{11}+ka_{21})a_{22}-(a_{12}+ka_{22})a_{21}$

$$=a_{11}a_{22}+ka_{21}a_{22}-a_{12}a_{21}-ka_{22}a_{21}=a_{11}a_{22}-a_{12}a_{21}$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$ □

下面考察 3 阶行列式.由定义得

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}\begin{vmatrix} a_{31} & a_{33} \\ a_{21} & a_{23} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \qquad (*)$$

在上述 3 阶行列式中，划去第 i 行第 j 列后所剩下的 2 阶行列式称为**元素 a_{ij} 的余子式**，记为 M_{ij}. 例如

$$M_{11}=\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix},\quad M_{12}=\begin{vmatrix}a_{21}&a_{23}\\a_{31}&a_{33}\end{vmatrix},\quad M_{13}=\begin{vmatrix}a_{21}&a_{22}\\a_{31}&a_{32}\end{vmatrix},$$

$$M_{23}=\begin{vmatrix}a_{11}&a_{12}\\a_{31}&a_{32}\end{vmatrix},\quad M_{31}=\begin{vmatrix}a_{12}&a_{13}\\a_{22}&a_{23}\end{vmatrix},\quad \text{等等.}$$

再令

$$A_{ij}=(-1)^{i+j}M_{ij},$$

称之为**元素 a_{ij} 的代数余子式**，例如

$$A_{11}=M_{11},\quad A_{12}=-M_{12},\quad A_{13}=M_{13},\quad A_{23}=-M_{23},\quad A_{31}=M_{31},$$

等等. 由上面的符号，(*) 式可写为

$$\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}=a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}. \tag{0.8}$$

同样地

$$\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}$$

$$=-a_{21}(-a_{13}a_{32}+a_{12}a_{33})+a_{22}(a_{11}a_{33}-a_{13}a_{31})-a_{23}(-a_{12}a_{31}+a_{11}a_{32})$$

$$=-a_{21}\begin{vmatrix}a_{12}&a_{13}\\a_{32}&a_{33}\end{vmatrix}+a_{22}\begin{vmatrix}a_{11}&a_{13}\\a_{31}&a_{33}\end{vmatrix}-a_{23}\begin{vmatrix}a_{11}&a_{12}\\a_{31}&a_{32}\end{vmatrix}$$

$$=-a_{21}M_{21}+a_{22}M_{22}-a_{23}M_{23}$$

$$=a_{21}A_{21}+a_{22}A_{22}+a_{23}A_{23}. \tag{0.9}$$

类似地，可得到

$$\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}=a_{31}A_{31}+a_{32}A_{32}+a_{33}A_{33}. \tag{0.10}$$

上面的(0.8)，(0.9)，(0.10)式分别称之为 3 阶行列式对其第一、二、三行的展开公式. 同样也有 3 阶行列式对其第一、二、三列的展开公式. 即

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31}$$

$$= a_{12}A_{12} + a_{22}A_{22} + a_{32}A_{32}$$

$$= a_{13}A_{13} + a_{23}A_{23} + a_{33}A_{33}.$$

这样,我们便证明了下面的定理.

定理 0.1 3 阶行列式等于它的任一行(或列)元素与该元素的代数余子式乘积之和.

当然,此定理对 2 阶行列式也成立.参见习题.

根据前面定理和已经证明的关于 2 阶行列式的性质,3 阶行列式也有同样的性质.

性质 0.1′ 行列互换,行列式的值不变.即

$$\begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

证明 等号左端的行列式按第 1 列展开后利用性质 0.1 可得

$$a_{11}\begin{vmatrix} a_{22} & a_{32} \\ a_{23} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{31} \\ a_{23} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{31} \\ a_{22} & a_{32} \end{vmatrix}$$

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

这恰为等号右端的行列式按第 1 行展开后的结果.

由此下面关于 3 阶行列式的性质也只对行来证明. □

性质 0.2′ 两行(列)互换,行列式的值变号.

证明 只考虑行列式的前 2 行互换的情形,其他情形类似.本性质即为

$$\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = -\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

把等号左端的行列式按第 3 行展开再利用性质 0.2 可得

$$a_{31}\begin{vmatrix} a_{22} & a_{23} \\ a_{12} & a_{13} \end{vmatrix} - a_{32}\begin{vmatrix} a_{21} & a_{23} \\ a_{11} & a_{13} \end{vmatrix} + a_{33}\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}$$

$$= -\left(a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \right)$$

$$=-\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$ □

类似地,可证明下面性质:

性质 0.3′ 若 3 阶行列式某行(列)各个元素分成两个元素之和,则该行列式可关于该行(列)拆开成两个行列式之和,拆开时其他各行(列)均保持不变.

性质 0.4′ 行列式某一行(列)的公因子 k 可以提到行列式外面.特别地,若行列式有一行(列)为零(即此行(列)元素全为 0),则行列式的值为 0.

性质 0.5′ 把一行(列)的倍数加到另一行(列)上,行列式的值不变.

例 0.2 证明

$$\begin{vmatrix} a_1+b_1 & b_1+c_1 & c_1+a_1 \\ a_2+b_2 & b_2+c_2 & c_2+a_2 \\ a_3+b_3 & b_3+c_3 & c_3+a_3 \end{vmatrix} = 2\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

证明 左端 $=\begin{vmatrix} a_1 & b_1+c_1 & c_1+a_1 \\ a_2 & b_2+c_2 & c_2+a_2 \\ a_3 & b_3+c_3 & c_3+a_3 \end{vmatrix}+\begin{vmatrix} b_1 & b_1+c_1 & c_1+a_1 \\ b_2 & b_2+c_2 & c_2+a_2 \\ b_3 & b_3+c_3 & c_3+a_3 \end{vmatrix}$

$$=\begin{vmatrix} a_1 & b_1+c_1 & c_1 \\ a_2 & b_2+c_2 & c_2 \\ a_3 & b_3+c_3 & c_3 \end{vmatrix}+\begin{vmatrix} b_1 & c_1 & c_1+a_1 \\ b_2 & c_2 & c_2+a_2 \\ b_3 & c_3 & c_3+a_3 \end{vmatrix}$$

$$=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}+\begin{vmatrix} b_1 & c_1 & a_1 \\ b_2 & c_2 & a_2 \\ b_3 & c_3 & a_3 \end{vmatrix}$$

$$=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}-\begin{vmatrix} c_1 & b_1 & a_1 \\ c_2 & b_2 & a_2 \\ c_3 & b_3 & a_3 \end{vmatrix}=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}+\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

=右端. □

例 0.3 证明

$$\begin{vmatrix} a_1 & 0 & 0 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}.$$

证明 把左端行列式按第一行展开即得. □

由上面例 0.3 可知,在计算 3 阶行列式时,我们可以利用行列式性质把某一行(或列)的 3 个元素中的 2 个变成 0. 然后再按此行(或列)展开就化为计算一个 2 阶行列式了.

例 0.4 计算下列行列式:

(1) $\begin{vmatrix} 1 & -1 & 2 \\ 3 & 2 & 1 \\ 0 & 1 & 4 \end{vmatrix}$; (2) $\begin{vmatrix} x & y & x+y \\ y & x+y & x \\ x+y & x & y \end{vmatrix}$; (3) $\begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{vmatrix}$.

解 (1) $\begin{vmatrix} 1 & -1 & 2 \\ 3 & 2 & 1 \\ 0 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 2 \\ 0 & 5 & -5 \\ 0 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 5 & -5 \\ 1 & 4 \end{vmatrix} = 25.$

(2) $$\begin{vmatrix} x & y & x+y \\ y & x+y & x \\ x+y & x & y \end{vmatrix} = \begin{vmatrix} 2x+2y & 2x+2y & 2x+2y \\ y & x+y & x \\ x+y & x & y \end{vmatrix} = 2(x+y)\begin{vmatrix} 1 & 1 & 1 \\ y & x+y & x \\ x+y & x & y \end{vmatrix}$$

$$= 2(x+y)\begin{vmatrix} 1 & 0 & 0 \\ y & x & x-y \\ x+y & -y & -x \end{vmatrix} = 2(x+y)\begin{vmatrix} x & x-y \\ -y & -x \end{vmatrix}$$

$$= -2(x+y)(x^2-xy+y^2) = -2(x^3+y^3).$$

(3) $$\begin{vmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ a_1 & a_2-a_1 & a_3-a_1 \\ a_1^2 & a_2^2-a_1^2 & a_3^2-a_1^2 \end{vmatrix}$$

$$= \begin{vmatrix} a_2-a_1 & a_3-a_1 \\ (a_2-a_1)(a_2+a_1) & (a_3-a_1)(a_3+a_1) \end{vmatrix}$$

$$= (a_2-a_1)(a_3-a_1)\begin{vmatrix} 1 & 1 \\ a_2+a_1 & a_3+a_1 \end{vmatrix}$$

$$= (a_2-a_1)(a_3-a_1)(a_3-a_2).$$

§0.3 求和号与连乘符号

数的加法和乘法这两种代数运算是代数学最基本的立足点(减法和除法分别是加法和乘法的逆运算,实际上可以归入加法和乘法).所以我们要不断地处理许多数的加法、乘法,为了把它们表达得清楚简明,我们引入下面的求和号与乘积号.

对于 n 个数 $a_1,a_2,\cdots,a_n$,记

$$a_1+a_2+\cdots+a_n=\sum_{i=1}^{n}a_i,$$

$$a_1a_2\cdots a_n=\prod_{i=1}^{n}a_i,$$

即 $\sum$ 和 $\prod$ 分别表示多个数连加与连乘.在上面两个式子中,求和号 $\sum$ 和乘积号 $\prod$ 后面写一个带变动下角标 i 的记号 a_i,在 $\sum$ 和 $\prod$ 的下上各加上"$i=1$"和"n"表示变动下角标 i 取从 1 开始到 n 的所有自然数,它恰好是等式左端连加或连乘的数值.当然,上面两式也可以写成

$$\sum_{i=1}^{n}a_i=\sum_{1\leqslant i\leqslant n}a_i,\quad \prod_{i=1}^{n}a_i=\prod_{1\leqslant i\leqslant n}a_i.$$

例如

$$\sum_{i=1}^{n}i=1+2+\cdots+n=\frac{1}{2}n(n+1),$$

而

$$\prod_{i=1}^{n}i=1\cdot 2\cdot\cdots\cdot n=n!.$$

我们知道

$$ka_1+ka_2+\cdots+ka_n=k(a_1+a_2+\cdots+a_n),$$

上面这个式子用求和号表示就是

$$\sum_{i=1}^{n}ka_i=k\sum_{i=1}^{n}a_i,$$

即求和号中每项的公因子可以提到求和号外面.又由加法的交换律和结合律有

$$(a_1+b_1)+(a_2+b_2)+\cdots+(a_n+b_n)=(a_1+a_2+\cdots+a_n)+(b_1+b_2+\cdots+b_n),$$

此式用求和号表示就是

$$\sum_{i=1}^{n}(a_i+b_i)=\sum_{i=1}^{n}a_i+\sum_{i=1}^{n}b_i,$$

所以若求和号中每一项都是两个数之和,则求和号可分为两个求和号之和,每个求和号的项都是原求和号中一项的其中一个数.

现在考察 mn 个数 $a_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$.把它们排成一个 m 行 n 列的数表如下:

$$\begin{matrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{matrix}$$

记第 i 行这 n 个数的和为 A_i,$1\leqslant i\leqslant m$,即

$$A_i=a_{i1}+a_{i2}+\cdots+a_{in}=\sum_{j=1}^{n}a_{ij}.$$

再把 $A_1,A_2,\cdots,A_m$ 连加起来，它恰为上述这 mn 个数 a_{ij} 连加的值，即为

$$\sum_{i=1}^{m}A_i=\sum_{i=1}^{m}\Big(\sum_{j=1}^{n}a_{ij}\Big).$$

下面考察另一种记法.用 B_j 表示第 j 列这 m 个数 $a_{1j},a_{2j},\cdots,a_{mj}$ 的和，$1\leqslant j\leqslant n$，即

$$B_j=a_{1j}+a_{2j}+\cdots+a_{mj}=\sum_{i=1}^{m}a_{ij},$$

再把 $B_1,B_2,\cdots,B_n$ 连加起来，这同样是上面这 mn 个数 a_{ij} 连加的结果，即为

$$\sum_{j=1}^{n}B_j=\sum_{j=1}^{n}\Big(\sum_{i=1}^{m}a_{ij}\Big).$$

上面用两种方式把 mn 个数 a_{ij} 连加，其结果当然是一样的，即有

$$\sum_{i=1}^{m}\Big(\sum_{j=1}^{n}a_{ij}\Big)=\sum_{j=1}^{n}\Big(\sum_{i=1}^{m}a_{ij}\Big).$$

上式表明：当考察有限个数连加时，出现的二重和号可以交换次序.这是利用和号作运算时十分有用的一个法则，上面式子中，小括号一般也可以不写，即简记为

$$\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}=\sum_{j=1}^{n}\sum_{i=1}^{m}a_{ij}.$$

类似地

$$\prod_{i=1}^{m}\Big(\prod_{j=1}^{n}a_{ij}\Big)$$

表示先把第 i 行的 n 个数连乘后再把所得的 m 个数相乘，结果即为把上面这 mn 个数 a_{ij} 连乘的结果，而

$$\prod_{j=1}^{n}\Big(\prod_{i=1}^{m}a_{ij}\Big)$$

是先把第 j 列上的 m 个数连乘后再把所得的 n 个数相乘，也是上面这 mn 个数 a_{ij} 连乘的结果.所以

$$\prod_{i=1}^{m}\Big(\prod_{j=1}^{n}a_{ij}\Big)=\prod_{j=1}^{n}\Big(\prod_{i=1}^{m}a_{ij}\Big),$$

即有限个数连乘时，出现的二重乘积号也可以交换次序.

进一步地，显然有

$$\begin{aligned}\Big(\sum_{i=1}^{m}a_i\Big)\cdot\Big(\sum_{j=1}^{n}b_j\Big)&=a_1\sum_{j=1}^{n}b_j+a_2\sum_{j=1}^{n}b_j+\cdots+a_m\sum_{j=1}^{n}b_j\\&=\sum_{j=1}^{n}a_1b_j+\sum_{j=1}^{n}a_2b_j+\cdots+\sum_{j=1}^{n}a_mb_j\\&=\sum_{i=1}^{m}\sum_{j=1}^{n}a_ib_j.\end{aligned}$$

§0.4 数 域

数及其四则运算是代数学最原始的出发点.因此,为了后面理论的展开有一个坚实的基础,就需要把我们熟知的数的概念提高一步,也就是把我们后面研讨的每一个课题涉及哪些数做出明确的交代.下面如无特殊说明,出现的数都来自于某个数域.为此先给出数域这个概念.

定义 0.1 设 F 是复数集的一个子集,$0,1\in F$,且 F 对加、减、乘、除(除数不为 0)运算封闭,即对于任意 $a,b\in F$,都有 $a+b,a-b,ab\in F$,且当 $b\neq 0$ 时 $\frac{a}{b}\in F$,则称 F 是一个**数域**.

显然,有理数集、实数集和复数集都是数域,分别称为有理数域、实数域和复数域,并分别记为$\mathbb{Q}$,$\mathbb{R}$ 和 $\mathbb{C}$,而整数集 $\mathbb{Z}$ 不是数域.

例 0.5 记 $\mathbb{Q}(\sqrt{3})=\{a+b\sqrt{3}\mid a,b\in\mathbb{Q}\}$,则 $\mathbb{Q}(\sqrt{3})$ 是一个数域.显然,$0=0+0\sqrt{3}\in\mathbb{Q}(\sqrt{3})$,$1=1+0\sqrt{3}\in\mathbb{Q}(\sqrt{3})$.对于任意 $\alpha=a+b\sqrt{3}$, $\beta=c+d\sqrt{3}\in\mathbb{Q}(\sqrt{3})$,其中 $a,b,c,d\in\mathbb{Q}$,则显然

$$\alpha\pm\beta=(a\pm c)+(b\pm d)\sqrt{3}.$$

$$\alpha\cdot\beta=(ac+3bd)+(ad+bc)\sqrt{3}.$$

由于 $a,b,c,d\in\mathbb{Q}$,所以 $a\pm c,b\pm d,ac+3bd,ad+bc\in\mathbb{Q}$.所以 $\alpha\pm\beta,\alpha\beta\in\mathbb{Q}(\sqrt{3})$.若 $\beta=c+d\sqrt{3}\neq 0$,则一定有 $c-d\sqrt{3}\neq 0$,所以

$$\frac{\alpha}{\beta}=\frac{a+b\sqrt{3}}{c+d\sqrt{3}}=\frac{(a+b\sqrt{3})(c-d\sqrt{3})}{(c+d\sqrt{3})(c-d\sqrt{3})}=\frac{ac-3bd}{c^2-3d^2}+\frac{bc-ad}{c^2-3d^2}\sqrt{3}\in\mathbb{Q}(\sqrt{3}).$$

这便证出 $\mathbb{Q}(\sqrt{3})$是一个数域.

类似地,容易验证,对任意素数 p,$\mathbb{Q}(\sqrt{p})=\{a+b\sqrt{p}\mid a,b\in\mathbb{Q}\}$是数域.所以数域有无穷多个.

命题 0.1 任一数域都包含有理数域.

证明 设 F 为一数域,则 $0,1\in F$,从而 $2=1+1\in F,3=2+1\in F,\cdots,n=(n-1)+1\in F$,即 F 包含所有正整数.对任意负整数$-n$,$-n=0-n\in F$.从而 F 包含所有整数.对任一有理数$\frac{a}{b}$,a,b 为整数且 $b\neq 0$,这是整数 a 除以整数 b 的结果.所以 $\frac{a}{b}\in F$,即 F 包含所有有理数. □

上面证明了,对任一数域 F 都有 $F\supseteq\mathbb{Q}$.所以一个数域中一定有无穷多个元素.

习 题 零

1. 证明 2 阶行列式等于它的任一行(或列)元素与自己的代数余子式乘积之和.

2. 证明性质 0.3′,0.4′和 0.5′.

3. 证明在一行列式(2 阶或 3 阶)中,若有两行(或列)元素相等(或成比例),则行列式值为 0.

4. 计算下列行列式:

(1) $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix}$; (2) $\begin{vmatrix} 2 & -4 & -3 \\ 7 & 2 & 3 \\ 4 & -2 & 6 \end{vmatrix}$;

(3) $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix}$; (4) $\begin{vmatrix} \lambda-2 & -3 & -2 \\ -1 & \lambda-8 & -2 \\ 2 & 14 & \lambda+3 \end{vmatrix}$.

5. 求下列和式的值;

(1) $\sum\limits_{i=2}^{n} i^2, \sum\limits_{i=1}^{n} (i+1)(i+2)$;

(2) $\sum\limits_{i=1}^{n} (-1)^i, \sum\limits_{i=5}^{15} (-1)^i i$.

6. 证明 $\sum\limits_{i=1}^{n} \dfrac{1}{i(i+1)} = \dfrac{n}{n+1}$.

7. 将 $(a+b)^n$ 的二项式展开公式用和号形式写出来.

8. 解下列三元一次方程组:

(1) $\begin{cases} x-y \quad =1, \\ 2x \quad +z=2, \\ 3x-y-z=0; \end{cases}$ (2) $\begin{cases} x+y-3z=-1, \\ 2x+y-2z=1, \\ x+y \quad +z=3. \end{cases}$

9. 证明二元一次方程组

$$\begin{cases} a_{11}x + a_{12}y = b_1, \\ a_{21}x + a_{22}y = b_2 \end{cases}$$

有唯一解的充分必要条件是

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0.$$

10. 证明 $\mathbb{Q}(\mathrm{i})=\{a+b\mathrm{i} \mid a,b\in\mathbb{Q}\}$ 为数域($\mathrm{i}=\sqrt{-1}$).

第一章 空间与向量

本章以向量代数为工具研究空间的平面与直线，首先介绍向量的基本概念、线性运算、内积、外积和混合积等向量代数知识，然后用它们来建立空间的平面和直线的方程，并讨论平面和直线相互间的位置（包括夹角和距离）关系.

§1.1 向量与空间直角坐标系

本节主要介绍向量的基本概念、向量的线性运算、空间直角坐标系、向量的坐标及点的坐标.

1.1.1 向量的基本概念

定义 1.1 一个既有大小又有方向的量叫做**向量**，用黑体字母表示，例如 $\boldsymbol{a},\boldsymbol{i},\boldsymbol{n},\boldsymbol{s},\boldsymbol{F}$ 等.

由于向量具有大小和方向两个要素，通常用几何中的有向线段表示向量，有向线段的长度表示向量的大小，有向线段的方向表示向量的方向，以 A 为始点、B 为终点的有向线段所表示的向量记作 $\overrightarrow{AB}$.

实际问题中的向量，有些与始点有关，有些与始点无关，与始点无关的向量叫做自由向量，后面我们只研究自由向量.

若向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的大小相等且方向相同，则称向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 相等，记做 $\boldsymbol{a}=\boldsymbol{b}$.也就是说，经过平移后能完全重合的两个向量称为相等的向量.若向量 $\boldsymbol{b}$ 的大小与 $\boldsymbol{a}$ 的大小相同，但方向相反，则 $\boldsymbol{b}$ 叫做 $\boldsymbol{a}$ 的反向量，记作 $\boldsymbol{b}=-\boldsymbol{a}$.

向量的大小叫做向量的长度（也称为模）.向量 $\boldsymbol{a}$ 的长度用 $|\boldsymbol{a}|$ 表示.

长度为 1 的向量叫做单位向量.

长度为零的向量叫做零向量，记作 $\boldsymbol{0}$.由于零向量是一个始点与终点重合的特殊向量，因

此它的方向可看作是任意的,使用时可随意选取.

对于向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$,将它们平移使始点重合.把 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的正向之间不大于 π 的角叫做向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角.

向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的平行和垂直分别用 $\boldsymbol{a}/\!/\boldsymbol{b}$ 和 $\boldsymbol{a}\perp\boldsymbol{b}$ 表示.

由于零向量的方向可以随意选取,因此零向量与任何向量都平行,也与任何向量都垂直.

因为将两个平行的向量平移到始点重合时,这两个向量在一条直线上,所以也把平行的向量称为共线向量.

如果三个向量都平行于同一个平面,那么将它们平移使之始点重合后,这三个向量就在一个平面上.因此,把平行于同一个平面的向量叫做共面向量.

请注意,上述向量的“共线”和“共面”是符合实际需要的数学概念,与直观上的共线(即在一条直线上)和共面(即在一个平面上)有所差异,其原因就在于所研究的向量是自由向量.

1.1.2 向量的线性运算及投影

定义 1.2 向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的和记作 $\boldsymbol{a}+\boldsymbol{b}$,规定 $\boldsymbol{a}+\boldsymbol{b}$ 是按“三角形法则”或“平行四边形法则”所确定的向量,即

(1) 三角形法则:

作 $\overrightarrow{OA}=\boldsymbol{a}$,$\overrightarrow{AC}=\boldsymbol{b}$,则 $\overrightarrow{OC}=\boldsymbol{a}+\boldsymbol{b}$(见图 1.1).

(2) 平行四边形法则:

设 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 不平行,作 $\overrightarrow{OA}=\boldsymbol{a}$,$\overrightarrow{OB}=\boldsymbol{b}$,以 $\overrightarrow{OA}$ 和 $\overrightarrow{OB}$ 为邻边作平行四边形 $OACB$,则 $\overrightarrow{OC}=\boldsymbol{a}+\boldsymbol{b}$(见图 1.2).

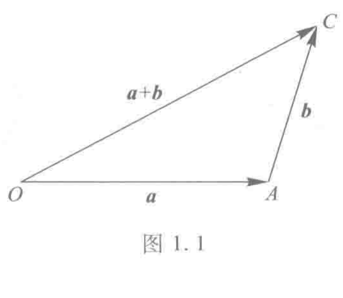


图 1.1

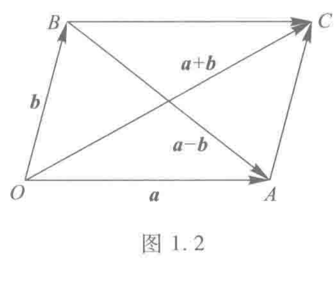


图 1.2

规定 $\boldsymbol{a}-\boldsymbol{b}=\boldsymbol{a}+(-\boldsymbol{b})$,它是一个以 $\boldsymbol{b}$ 的终点为始点,$\boldsymbol{a}$ 的终点为终点的向量(见图 1.2 中的 $\overrightarrow{BA}$).

定义 1.3 实数 λ 与向量 $\boldsymbol{a}$ 的乘积记作 $\lambda\boldsymbol{a}$,规定 $\lambda\boldsymbol{a}$ 是一个向量,它的长度为

$$|\lambda\boldsymbol{a}|=|\lambda||\boldsymbol{a}|,$$

它的方向当 $\lambda>0$ 时与 $\boldsymbol{a}$ 相同；当 $\lambda<0$ 时与 $\boldsymbol{a}$ 相反；当 $\lambda=0$ 时，因为 $\lambda\boldsymbol{a}=\boldsymbol{0}$，其方向任意.

向量的加法和向量与实数的乘法这两种运算统称为向量的线性运算.

设 λ 和 μ 是实数，$\boldsymbol{a},\boldsymbol{b}$ 和 $\boldsymbol{c}$ 是任意三个向量.由定义容易验证向量的线性运算具有下列八条性质：

(1) $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$;

(2) $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$;

(3) $\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{a}$;

(4) $\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}$;

(5) $\lambda(\mu\boldsymbol{a})=\mu(\lambda\boldsymbol{a})=(\lambda\mu)\boldsymbol{a}$;

(6) $1\boldsymbol{a}=\boldsymbol{a}$ 且 $(-1)\boldsymbol{a}=-\boldsymbol{a}$;

(7) $\lambda(\boldsymbol{a}+\boldsymbol{b})=\lambda\boldsymbol{a}+\lambda\boldsymbol{b}$;

(8) $(\lambda+\mu)\boldsymbol{a}=\lambda\boldsymbol{a}+\mu\boldsymbol{a}$.

由于向量的加法满足交换律和结合律，因此三个向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 的和可简记为 $\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}$.先以向量 $\boldsymbol{a}$ 的终点为始点作向量 $\boldsymbol{b}$，再以向量 $\boldsymbol{b}$ 的终点为始点作向量 $\boldsymbol{c}$，由三角形法则可知，$\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}$ 是一个以 $\boldsymbol{a}$ 的始点为始点、以 $\boldsymbol{c}$ 的终点为终点的向量.推广到一般情况，n 个向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n(n\geqslant 2)$ 之和为 $\boldsymbol{a}_1+\boldsymbol{a}_2+\cdots+\boldsymbol{a}_n$，设 $\boldsymbol{a}_i$ 的起点为 $\boldsymbol{a}_{i-1}$ 的终点，$2\leqslant i\leqslant n$，则它是一个以 $\boldsymbol{a}_1$ 的始点为始点，以 $\boldsymbol{a}_n$ 的终点为终点的向量.

当 $\boldsymbol{b}\neq\boldsymbol{0}$ 时，与 $\boldsymbol{b}$ 同方向的单位向量为 $\dfrac{\boldsymbol{b}}{|\boldsymbol{b}|}$.

设向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的夹角为 θ 且 $\boldsymbol{b}\neq\boldsymbol{0}$，把 $|\boldsymbol{a}|\cos\theta$ 称为向量 $\boldsymbol{a}$ 在 $\boldsymbol{b}$ 上的投影，记作 $(\boldsymbol{a})_{\boldsymbol{b}}$，即

$$(\boldsymbol{a})_{\boldsymbol{b}}=|\boldsymbol{a}|\cos\theta.$$

把 $(\boldsymbol{a})_{\boldsymbol{b}}\dfrac{\boldsymbol{b}}{|\boldsymbol{b}|}$ 叫做向量 $\boldsymbol{a}$ 在 $\boldsymbol{b}$ 上的投影向量(见图 1.3).

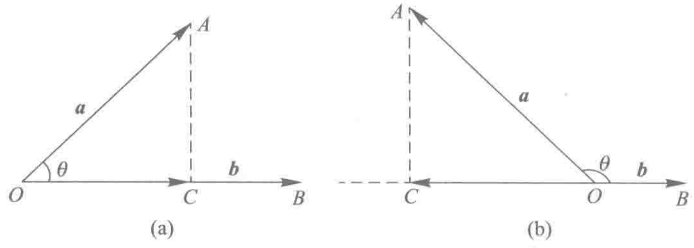


图 1.3

设 $\boldsymbol{c}\neq\boldsymbol{0}$，可以验证，向量的投影具有下列性质：

(1) 若 $\boldsymbol{a}=\boldsymbol{b}$，则 $(\boldsymbol{a})_{\boldsymbol{c}}=(\boldsymbol{b})_{\boldsymbol{c}}$;

(2) $\forall\lambda\in\mathbb{R}$，都有 $(\lambda\boldsymbol{a})_{\boldsymbol{c}}=\lambda(\boldsymbol{a})_{\boldsymbol{c}}$;

(3) $(\boldsymbol{a}+\boldsymbol{b})_{\boldsymbol{c}}=(\boldsymbol{a})_{\boldsymbol{c}}+(\boldsymbol{b})_{\boldsymbol{c}}$.

1.1.3 空间直角坐标系

在平面解析几何中，通过建立的平面直角坐标系，将平面上的点与坐标一一对应，就能用代数的方法来讨论平面上几何图形的性质.我们将上述思想推广，建立空间直角坐标系.

在空间选定一点 O 作为原点，过 O 点作三个两两垂直的数轴.这三个数轴统称为坐标轴，分别用 x 轴、y 轴和 z 轴表示，由原点 O 和这三个坐标轴所组成的系统叫做空间直角坐标系，简称"$Oxyz$ 坐标系".

在空间直角坐标系中，每两个坐标轴都决定一个平面，分别称为 Oxy 面、Oyz 面和 Ozx 面，统称为坐标面.

在 $Oxyz$ 坐标系中，分别与 x 轴、y 轴和 z 轴的正方向同向的单位向量叫做这个坐标系的基向量，用 $\boldsymbol{i}$,$\boldsymbol{j}$ 和 $\boldsymbol{k}$ 表示.

设 $\boldsymbol{p}$ 为空间中任一给定的向量，$\overrightarrow{OP}=\boldsymbol{p}$，过 P 点作三个平面分别垂直于 x 轴、y 轴和 z 轴，它们与这三个坐标轴的交点依次记为 A，B 和 C（见图 1.4）.

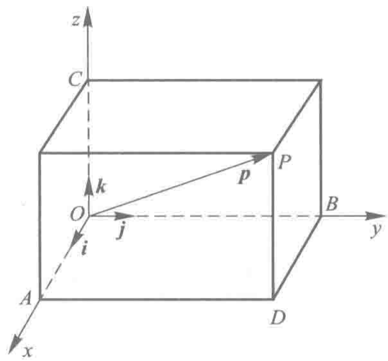


图 1.4

设向量$\overrightarrow{OP}$在三个坐标轴上的投影依次为 p_x,p_y,p_z，由 $PA\perp x$ 轴、$PB\perp y$ 轴、$PC\perp z$ 轴可知，$\overrightarrow{OP}$在这三个坐标轴上的投影向量依次为

$$\overrightarrow{OA}=p_x\boldsymbol{i},\quad \overrightarrow{OB}=p_y\boldsymbol{j},\quad \overrightarrow{OC}=p_z\boldsymbol{k}.$$

由于$\overrightarrow{OP}=\overrightarrow{OA}+\overrightarrow{AD}+\overrightarrow{DP}$，$\overrightarrow{AD}=\overrightarrow{OB}$，$\overrightarrow{DP}=\overrightarrow{OC}$，因此

$$\boldsymbol{p}=p_x\boldsymbol{i}+p_y\boldsymbol{j}+p_z\boldsymbol{k}.$$

不难看出，在 $Oxyz$ 坐标系中，向量 $\boldsymbol{p}$ 与有序数组(p_x,p_y,p_z)一一对应.

我们把 p_x,p_y,p_z 叫做向量 $\boldsymbol{p}$ 分别在 x 轴、y 轴和 z 轴上的坐标，把有序数组(p_x,p_y,p_z)叫做向量 $\boldsymbol{p}$ 在 $Oxyz$ 坐标系中的坐标，因此 $\boldsymbol{p}=p_x\boldsymbol{i}+p_y\boldsymbol{j}+p_z\boldsymbol{k}$ 可简记作

$$\boldsymbol{p}=(p_x,p_y,p_z).$$

下面我们用向量的坐标来表示向量的长度.

设非零向量 $\boldsymbol{p}=p_x\boldsymbol{i}+p_y\boldsymbol{j}+p_z\boldsymbol{k}$，$\overrightarrow{OP}=\boldsymbol{p}$（见图 1.4）.利用勾股定理不难得出

$$|\overrightarrow{OP}|^2=|\overrightarrow{OA}|^2+|\overrightarrow{OB}|^2+|\overrightarrow{OC}|^2,$$

于是，$\boldsymbol{p}$ 的长度为

$$|\boldsymbol{p}|=\sqrt{p_x^2+p_y^2+p_z^2}.$$

对于上面的向量 $\boldsymbol{p}$，它的方向可用与其同向的单位向量来表示，即

$$\frac{\boldsymbol{p}}{|\boldsymbol{p}|}=\frac{p_x}{|\boldsymbol{p}|}\boldsymbol{i}+\frac{p_y}{|\boldsymbol{p}|}\boldsymbol{j}+\frac{p_z}{|\boldsymbol{p}|}\boldsymbol{k}.$$

由向量的投影的定义可得

$$\frac{\boldsymbol{p}}{|\boldsymbol{p}|}=\cos\alpha\boldsymbol{i}+\cos\beta\boldsymbol{j}+\cos\gamma\boldsymbol{k},$$

其中 α,β,γ 分别是向量 $\boldsymbol{p}$ 与 x 轴、y 轴和 z 轴正向的夹角.

因为用 α,β,γ 也能确定向量 $\boldsymbol{p}$ 的方向，所以我们把它们叫做向量 $\boldsymbol{p}$ 的**方向角**，它们的余弦 $\cos\alpha,\cos\beta,\cos\gamma$ 叫做向量 $\boldsymbol{p}$ 的**方向余弦**，其中

$$\cos\alpha=\frac{p_x}{\sqrt{p_x^2+p_y^2+p_z^2}},$$

$$\cos\beta=\frac{p_y}{\sqrt{p_x^2+p_y^2+p_z^2}},$$

$$\cos\gamma=\frac{p_z}{\sqrt{p_x^2+p_y^2+p_z^2}},$$

显然，

$$\cos^2\alpha+\cos^2\beta+\cos^2\gamma=1.$$

这表明非零向量的三个方向角（或方向余弦）不是独立的，它们必须满足上述关系式.

下面我们来研究空间中点的坐标及两点之间的距离.

对于空间中任一给定点 P，由于点 P 与向量 $\overrightarrow{OP}$ 一一对应，把向量 $\overrightarrow{OP}$ 的坐标 p_x,p_y,p_z 依次称为点 P 的横坐标、纵坐标和竖坐标，并用 $P(p_x,p_y,p_z)$ 表示.

对于始点为 $P(p_x,p_y,p_z)$、终点为 $Q(q_x,q_y,q_z)$ 的向量 $\overrightarrow{PQ}$，由

$$\overrightarrow{OP}=p_x\boldsymbol{i}+p_y\boldsymbol{j}+p_z\boldsymbol{k},\quad \overrightarrow{OQ}=q_x\boldsymbol{i}+q_y\boldsymbol{j}+q_z\boldsymbol{k},$$

可得

$$\overrightarrow{PQ}=\overrightarrow{OQ}-\overrightarrow{OP}=(q_x-p_x)\boldsymbol{i}+(q_y-p_y)\boldsymbol{j}+(q_z-p_z)\boldsymbol{k}.$$

由此看出，空间中任一向量的坐标就是其终点的坐标与始点的坐标之差.

把两点 $P(p_x,p_y,p_z)$ 和 $Q(q_x,q_y,q_z)$ 之间的距离记作 $d(P,Q)$，则

$$d(P,Q)=|\overrightarrow{PQ}|=\sqrt{(q_x-p_x)^2+(q_y-p_y)^2+(q_z-p_z)^2}.$$

对于特殊的点，其坐标也有相应的特点.

不难得出，点 P 在坐标轴或坐标面上 $\Leftrightarrow$ 点 P 的三个坐标中至少有一个为零.具体情况有 7 种，下面列出 3 种，其余 4 种可类似得出：

(1) 点 P 为坐标原点 $\Leftrightarrow P$ 的三个坐标均为零；

(2) 点 P 在 x 轴上 $\Leftrightarrow P$ 的纵坐标和竖坐标均为零；

(3) 点 P 在 Oxy 面上 $\Leftrightarrow P$ 的竖坐标为零.

例 1.1 求点 $P(1,2,0)$ 到点 $Q(2,1,\sqrt{2})$ 的距离 $d(P,Q)$，向量 $\overrightarrow{PQ}$ 的长度、方向余弦、方向角，以及与其同向的单位向量.

解

$$\overrightarrow{PQ}=(2-1)\boldsymbol{i}+(1-2)\boldsymbol{j}+(\sqrt{2}-0)\boldsymbol{k}=\boldsymbol{i}-\boldsymbol{j}+\sqrt{2}\boldsymbol{k},$$

$$|\overrightarrow{PQ}|=\sqrt{1^2+(-1)^2+(\sqrt{2})^2}=2,$$

$$d(P,Q)=|\overrightarrow{PQ}|=2,$$

$$\cos\alpha=\frac{1}{2},\quad \cos\beta=-\frac{1}{2},\quad \cos\gamma=\frac{\sqrt{2}}{2},$$

$$\alpha=\frac{\pi}{3},\quad \beta=\frac{2\pi}{3},\quad \gamma=\frac{\pi}{4},$$

与$\overrightarrow{PQ}$同向的单位向量为

$$\frac{\overrightarrow{PQ}}{|\overrightarrow{PQ}|}=\frac{1}{2}\boldsymbol{i}-\frac{1}{2}\boldsymbol{j}+\frac{\sqrt{2}}{2}\boldsymbol{k}.$$

空间中全体向量的集合叫做三维几何空间，在 $Oxyz$ 坐标系中，它可以用从原点出发的所有向量的集合来表示.

§ 1.2 向量的内积、外积与混合积

在上一节我们介绍了向量的线性运算.在本节将介绍向量的内积、外积以及混合积这三种运算，研究它们的几何意义、性质和在空间直角坐标系中的坐标表达式，并由此研究向量间的关系.

1.2.1 内积

在物理中，物体在与水平线成 θ 角的力 $\boldsymbol{F}$ 的作用下，若物体仅有水平位移 $\boldsymbol{s}$，则力 $\boldsymbol{F}$ 对物体所做的功为

$$W=|\boldsymbol{F}||\boldsymbol{s}|\cos\theta.$$

我们将这个求功的运算抽象成如下的概念：

定义 1.4 两个向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 内积（也叫数量积或点乘积）是一个数，记作 $\boldsymbol{a}\cdot\boldsymbol{b}$，读作“$\boldsymbol{a}$ 点乘 $\boldsymbol{b}$”，规定

$$\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}||\boldsymbol{b}|\cos\theta,$$

其中 θ 为 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角.

根据定义 1.4，可知

$$W=\boldsymbol{F}\cdot\boldsymbol{s},$$

这是内积的物理定义.

内积与向量的投影有着密切的联系:

$$\boldsymbol{a}\cdot\boldsymbol{b}=\begin{cases}|\boldsymbol{b}|(\boldsymbol{a})_{\boldsymbol{b}}, & 当\ \boldsymbol{b}\neq\boldsymbol{0}\ 时;\\ |\boldsymbol{a}|(\boldsymbol{b})_{\boldsymbol{a}}, & 当\ \boldsymbol{a}\neq\boldsymbol{0}\ 时.\end{cases}$$

内积具有下列性质(其中 λ 为实数):

(1) $\boldsymbol{a}\cdot\boldsymbol{a}\geqslant 0,\boldsymbol{a}\cdot\boldsymbol{a}=0\Leftrightarrow\boldsymbol{a}=\boldsymbol{0}$;

(2) $\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{b}\cdot\boldsymbol{a}$;

(3) $(\lambda\boldsymbol{a})\cdot\boldsymbol{b}=\boldsymbol{a}\cdot(\lambda\boldsymbol{b})=\lambda(\boldsymbol{a}\cdot\boldsymbol{b})$;

(4) $(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{c}=\boldsymbol{a}\cdot\boldsymbol{c}+\boldsymbol{b}\cdot\boldsymbol{c}$.

由内积的定义容易验证(1)、(2)和(3)成立.下面证明(4)也成立.

证明 当 $\boldsymbol{c}=\boldsymbol{0}$ 时,(4)显然成立.当 $\boldsymbol{c}\neq\boldsymbol{0}$ 时,由向量的投影的性质得出

$$\begin{aligned}(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{c}&=|\boldsymbol{c}|(\boldsymbol{a}+\boldsymbol{b})_{\boldsymbol{c}}=|\boldsymbol{c}|[(\boldsymbol{a})_{\boldsymbol{c}}+(\boldsymbol{b})_{\boldsymbol{c}}]\\&=|\boldsymbol{c}|(\boldsymbol{a})_{\boldsymbol{c}}+|\boldsymbol{c}|(\boldsymbol{b})_{\boldsymbol{c}}=\boldsymbol{a}\cdot\boldsymbol{c}+\boldsymbol{b}\cdot\boldsymbol{c},\end{aligned}$$

因此(4)成立.

下面我们利用内积的性质,导出内积在空间直角坐标系中的坐标表达式.

设 $\boldsymbol{a}=a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k},\boldsymbol{b}=b_x\boldsymbol{i}+b_y\boldsymbol{j}+b_z\boldsymbol{k}$,由

$$\boldsymbol{i}\cdot\boldsymbol{j}=\boldsymbol{j}\cdot\boldsymbol{k}=\boldsymbol{k}\cdot\boldsymbol{i}=0,\quad \boldsymbol{i}\cdot\boldsymbol{i}=\boldsymbol{j}\cdot\boldsymbol{j}=\boldsymbol{k}\cdot\boldsymbol{k}=1,$$

可得

$$\begin{aligned}\boldsymbol{a}\cdot\boldsymbol{b}&=(a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k})\cdot(b_x\boldsymbol{i}+b_y\boldsymbol{j}+b_z\boldsymbol{k})\\&=a_xb_x+a_yb_y+a_zb_z,\end{aligned}$$

即两个向量的内积等于它们的坐标的乘积之和.

设向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角为 θ,当 $|\boldsymbol{a}||\boldsymbol{b}|\neq 0$ 时,由 $\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}||\boldsymbol{b}|\cos\theta$ 可得,

$$\cos\theta=\frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|}=\frac{a_xb_x+a_yb_y+a_zb_z}{\sqrt{a_x^2+a_y^2+a_z^2}\sqrt{b_x^2+b_y^2+b_z^2}}.$$

由此不难求出向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的夹角 θ,以及 $\boldsymbol{a}$ 在 $\boldsymbol{b}$ 上的投影与投影向量.

特别地,两个向量垂直的充要条件是它们的内积为零,即 $\boldsymbol{a}\perp\boldsymbol{b}\Leftrightarrow\boldsymbol{a}\cdot\boldsymbol{b}=0$. 两个向量垂直又称为两个向量正交.注意,我们规定零向量与任何向量垂直,也与任何向量共线.

例 1.2 设 $\boldsymbol{a}=2\boldsymbol{i}+\boldsymbol{j}-2\boldsymbol{k},\boldsymbol{b}=\boldsymbol{i}+\boldsymbol{j}$,求 $\boldsymbol{a}\cdot\boldsymbol{b}$,$\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角 θ,以及向量 $\boldsymbol{a}$ 在 $\boldsymbol{b}$ 上的投影和投影向量.

解

$$\boldsymbol{a}\cdot\boldsymbol{b}=2\times1+1\times1+(-2)\times0=3.$$

$$|\boldsymbol{a}|=\sqrt{2^2+1^2+(-2)^2}=3,\text{同理}\ |\boldsymbol{b}|=\sqrt{2},$$

$$\cos\theta=\frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|}=\frac{3}{3\times\sqrt{2}}=\frac{\sqrt{2}}{2},\quad \theta=\frac{\pi}{4}.$$

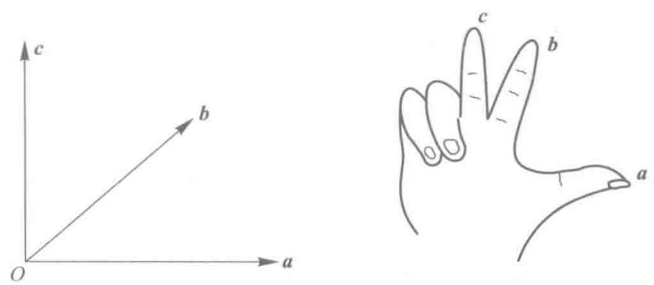
$$(\boldsymbol{a})_{b}=|\boldsymbol{a}|\cos\theta=\frac{3}{2}\sqrt{2},$$

$\boldsymbol{a}$ 在 $\boldsymbol{b}$ 上的投影向量为

$$(\boldsymbol{a})_{b}\frac{\boldsymbol{b}}{|\boldsymbol{b}|}=\frac{3}{2}\boldsymbol{i}+\frac{3}{2}\boldsymbol{j}.$$

1.2.2　外积

给定空间中不共面的三个向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$，假设它们的始点都是 O.如图 1.5 所示，如果此时 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 的位置关系和右手的拇指、食指、中指的位置关系相同，我们就称 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 构成右手系.



(a)　　(b)

图 1.5

定义 1.5　两个向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的外积（也叫矢量积或叉乘积）是一个向量，记作 $\boldsymbol{a}\times\boldsymbol{b}$，读作“$\boldsymbol{a}$ 叉乘 $\boldsymbol{b}$”，规定：

(1) 它的长度为 $|\boldsymbol{a}\times\boldsymbol{b}|=|\boldsymbol{a}||\boldsymbol{b}|\sin\theta$，其中 θ 是 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角；

(2) 它的方向与 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 都垂直，且按 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{a}\times\boldsymbol{b}$ 的次序构成右手系.

根据定义 1.5，可知当 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 平行时，有 $\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{0}$.

外积的几何意义是：当 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 不平行时，

(1) $|\boldsymbol{a}\times\boldsymbol{b}|$ 表示以 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 为邻边的平行四边形的面积；

(2) 向量 $\boldsymbol{a}\times\boldsymbol{b}$ 是由 $\boldsymbol{a}$ 在 $\boldsymbol{b}$ 的法平面 α（即 $\boldsymbol{b}\perp\alpha$）上的投影向量 $\boldsymbol{a}'$ 绕着 $\boldsymbol{b}$，按右手法则（即 $\boldsymbol{a}',\boldsymbol{b},\boldsymbol{a}\times\boldsymbol{b}$ 构成右手系）旋转 90°，再乘以 $|\boldsymbol{b}|$ 而得到的，如图 1.6.

外积具有下列性质：

(1) $\boldsymbol{a}\times\boldsymbol{a}=\boldsymbol{0}$；

(2) $\boldsymbol{a}\times\boldsymbol{b}=-\boldsymbol{b}\times\boldsymbol{a}$（反对称性）；

(3) $(\lambda\boldsymbol{a})\times\boldsymbol{b}=\boldsymbol{a}\times(\lambda\boldsymbol{b})=\lambda(\boldsymbol{a}\times\boldsymbol{b})$；

(4) $(\boldsymbol{a}+\boldsymbol{b})\times\boldsymbol{c}=\boldsymbol{a}\times\boldsymbol{c}+\boldsymbol{b}\times\boldsymbol{c}$.

根据定义 1.5 不难证明前三个性质，下面说明(4)也成立.当 $\boldsymbol{c}=\boldsymbol{0}$ 时显然成立；当 $\boldsymbol{c}\neq\boldsymbol{0}$ 时，根据外积的几何意义，向量 $(\boldsymbol{a}+\boldsymbol{b})\times\boldsymbol{c}$、$\boldsymbol{a}\times\boldsymbol{c}$ 和 $\boldsymbol{b}\times\boldsymbol{c}$ 分别是向量 $\boldsymbol{a}+\boldsymbol{b}$、$\boldsymbol{a}$ 和 $\boldsymbol{b}$ 在 $\boldsymbol{c}$ 的法平面 β 上的投影向量 $(\boldsymbol{a}+\boldsymbol{b})'$、$\boldsymbol{a}'$ 和 $\boldsymbol{b}'$，经过绕着 $\boldsymbol{c}$ 按右手法则旋转 90°，再乘以 $|\boldsymbol{c}|$ 而得到的.

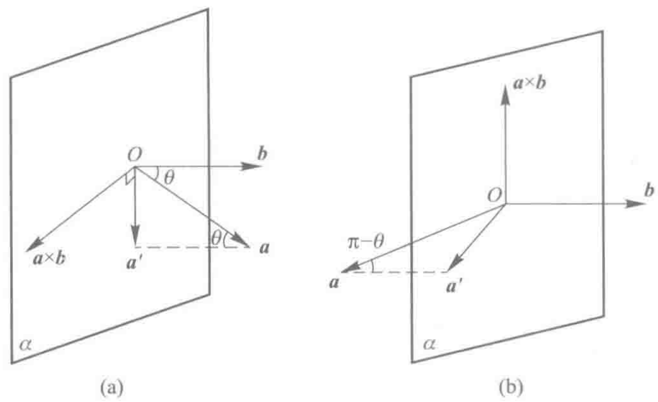


图 1.6

由于向量 $\boldsymbol{a}+\boldsymbol{b}$、$\boldsymbol{a}$ 和 $\boldsymbol{b}$ 满足三角形法则,而投影保持这种首尾相连的关系,所以

$$(\boldsymbol{a}+\boldsymbol{b})'=\boldsymbol{a}'+\boldsymbol{b}'.$$

根据右手法则不难验证向量 $(\boldsymbol{a}+\boldsymbol{b})'$、$\boldsymbol{a}'$ 和 $\boldsymbol{b}'$ 在法平面 β 上的旋转方向相同,故旋转后仍保持向量加法关系,再由向量线性运算的性质(7)可得(4)式成立.

下面推导外积在空间直角坐标系中的坐标表达式.为方便起见,本书中的基向量 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 均构成右手系.

设 $\boldsymbol{a}=a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k}$, $\boldsymbol{b}=b_x\boldsymbol{i}+b_y\boldsymbol{j}+b_z\boldsymbol{k}$,由

$$\boldsymbol{i}\times\boldsymbol{j}=-\boldsymbol{j}\times\boldsymbol{i}=\boldsymbol{k},\quad \boldsymbol{j}\times\boldsymbol{k}=-\boldsymbol{k}\times\boldsymbol{j}=\boldsymbol{i},$$
$$\boldsymbol{k}\times\boldsymbol{i}=-\boldsymbol{i}\times\boldsymbol{k}=\boldsymbol{j},\quad \boldsymbol{i}\times\boldsymbol{i}=\boldsymbol{j}\times\boldsymbol{j}=\boldsymbol{k}\times\boldsymbol{k}=\boldsymbol{0}$$

及外积的运算性质得出,

$$\begin{aligned}\boldsymbol{a}\times\boldsymbol{b}&=(a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k})\times(b_x\boldsymbol{i}+b_y\boldsymbol{j}+b_z\boldsymbol{k})\\&=a_xb_x(\boldsymbol{i}\times\boldsymbol{i})+a_yb_y(\boldsymbol{j}\times\boldsymbol{j})+a_zb_z(\boldsymbol{k}\times\boldsymbol{k})\\&\quad+a_xb_y(\boldsymbol{i}\times\boldsymbol{j})+a_xb_z(\boldsymbol{i}\times\boldsymbol{k})+a_yb_x(\boldsymbol{j}\times\boldsymbol{i})\\&\quad+a_yb_z(\boldsymbol{j}\times\boldsymbol{k})+a_zb_x(\boldsymbol{k}\times\boldsymbol{i})+a_zb_y(\boldsymbol{k}\times\boldsymbol{j})\\&=(a_yb_z-a_zb_y)\boldsymbol{i}-(a_xb_z-a_zb_x)\boldsymbol{j}+(a_xb_y-a_yb_x)\boldsymbol{k}.\end{aligned}$$

利用第零章中定义的行列式,有

$$\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix}a_y&a_z\\b_y&b_z\end{vmatrix}\boldsymbol{i}-\begin{vmatrix}a_x&a_z\\b_x&b_z\end{vmatrix}\boldsymbol{j}+\begin{vmatrix}a_x&a_y\\b_x&b_y\end{vmatrix}\boldsymbol{k}=\begin{vmatrix}\boldsymbol{i}&\boldsymbol{j}&\boldsymbol{k}\\a_x&a_y&a_z\\b_x&b_y&b_z\end{vmatrix}.$$

注意 最后一步只是借用三阶行列式的形式便于记忆,其中 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 是向量.

例 1.3 设向量 $\boldsymbol{a}=-\boldsymbol{i}+\boldsymbol{j}+\boldsymbol{k}$, $\boldsymbol{b}=2\boldsymbol{i}-\boldsymbol{j}+\boldsymbol{k}$,求 $\boldsymbol{a}\times\boldsymbol{b}$ 及以 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 为邻边平行四边形的面积 S.

解
$$\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix}\boldsymbol{i}&\boldsymbol{j}&\boldsymbol{k}\\-1&1&1\\2&-1&1\end{vmatrix}=2\boldsymbol{i}+3\boldsymbol{j}-\boldsymbol{k},$$
$$S=|\boldsymbol{a}\times\boldsymbol{b}|=\sqrt{2^2+3^2+(-1)^2}=\sqrt{14}.$$

1.2.3 混合积

定义 1.6 三个向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 的混合积是一个数,记作$(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$,规定
$$(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}.$$

混合积的几何意义如下:

以三个非零向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 为棱作一个平行六面体,其底面积为$|\boldsymbol{a}\times\boldsymbol{b}|$,高为$|\boldsymbol{c}||\cos\theta|$(其中 θ 是 $\boldsymbol{c}$ 与 $\boldsymbol{a}\times\boldsymbol{b}$ 的夹角).于是,该平行六面体的体积为
$$V=|\boldsymbol{a}\times\boldsymbol{b}||\boldsymbol{c}||\cos\theta|=|(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}|=|(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})|.$$

由此可见,三个向量的混合积的绝对值表示以这三个向量为棱的平行六面体的体积.

不难看出,若三个向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 依该次序符合右手法则,则 $V=(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$(图 1.7(a));否则,$V=-(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$(图 1.7(b)).

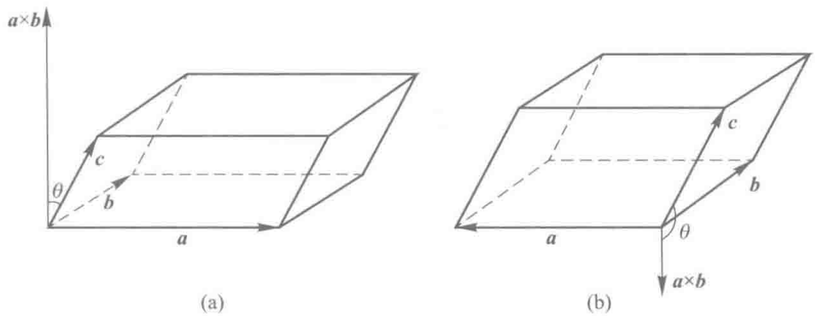


图 1.7

下面在空间直角坐标系中建立混合积的坐标表达式.

设 $\boldsymbol{a}=a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k},\boldsymbol{b}=b_x\boldsymbol{i}+b_y\boldsymbol{j}+b_z\boldsymbol{k},\boldsymbol{c}=c_x\boldsymbol{i}+c_y\boldsymbol{j}+c_z\boldsymbol{k}$,由
$$\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix}a_y&a_z\\b_y&b_z\end{vmatrix}\boldsymbol{i}-\begin{vmatrix}a_x&a_z\\b_x&b_z\end{vmatrix}\boldsymbol{j}+\begin{vmatrix}a_x&a_y\\b_x&b_y\end{vmatrix}\boldsymbol{k},$$
得出
$$\begin{aligned}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})&=(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}\\&=\begin{vmatrix}a_y&a_z\\b_y&b_z\end{vmatrix}c_x-\begin{vmatrix}a_x&a_z\\b_x&b_z\end{vmatrix}c_y+\begin{vmatrix}a_x&a_y\\b_x&b_y\end{vmatrix}c_z\end{aligned}$$

$$=\begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}=\begin{vmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{vmatrix}.$$

不难验证,混合积具有下列性质:

(1) $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=(\boldsymbol{c},\boldsymbol{a},\boldsymbol{b})=(\boldsymbol{b},\boldsymbol{c},\boldsymbol{a})$;

(2) $(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=-(\boldsymbol{b},\boldsymbol{a},\boldsymbol{c})$;

(3) $(\boldsymbol{a}_1+\boldsymbol{a}_2,\boldsymbol{b},\boldsymbol{c})=(\boldsymbol{a}_1,\boldsymbol{b},\boldsymbol{c})+(\boldsymbol{a}_2,\boldsymbol{b},\boldsymbol{c})$;

(4) $(\lambda\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=\lambda(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$.

(5) 若 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 中有两个相等或有一个为零向量,则$(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=0$.

例 1.4 设 $\boldsymbol{a}=\boldsymbol{i}+\boldsymbol{j},\boldsymbol{b}=\boldsymbol{j}+\boldsymbol{k},\boldsymbol{c}=-\boldsymbol{k}-\boldsymbol{i}$,它们的始点同为原点 O,求以 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 为棱的平行六面体的体积 V.

解 因为

$$(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=\begin{vmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & -1 \end{vmatrix}=-2,$$

所以 $V=2$.

1.2.4 向量间的关系

下面利用向量代数研究向量间的关系,从而为在空间直角坐标系中建立平面方程和直线方程,研究平面与直线间的关系奠定基础.

当向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 中有一个为零向量时,$\boldsymbol{a}$ 与 $\boldsymbol{b}$ 既平行又垂直;当 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 中有一个为零向量时,$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 一定共面.下面对 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 都不是零向量的情况加以研究.

定理 1.1 设 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 都是非零向量,θ 为 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角,则

(1) $\boldsymbol{a}\perp\boldsymbol{b}\Leftrightarrow\boldsymbol{a}\cdot\boldsymbol{b}=0$;

(2) $\boldsymbol{a}/\!/\boldsymbol{b}\Leftrightarrow\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{0}$;

(3) $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 共面$\Leftrightarrow(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=0$.

证明 (1) 根据 $\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}||\boldsymbol{b}|\cos\theta$ 及 $|\boldsymbol{a}||\boldsymbol{b}|\neq0$,可得

$$\boldsymbol{a}\perp\boldsymbol{b}\Leftrightarrow\theta=\frac{\pi}{2}\Leftrightarrow\boldsymbol{a}\cdot\boldsymbol{b}=0.$$

(2) 根据 $|\boldsymbol{a}\times\boldsymbol{b}|=|\boldsymbol{a}||\boldsymbol{b}|\sin\theta$ 及 $|\boldsymbol{a}||\boldsymbol{b}|\neq0$,可得

$$\boldsymbol{a}/\!/\boldsymbol{b}\Leftrightarrow\theta=0(\text{或 }\pi)\Leftrightarrow\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{0}.$$

(3) 当 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 共面时,若 $\boldsymbol{a}/\!/\boldsymbol{b}$,则 $\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{0}$.于是$(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=0$.

若 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 不平行,因为 $\boldsymbol{a}\times\boldsymbol{b}$ 垂直于 $\boldsymbol{a},\boldsymbol{b}$ 所在的平面,且 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 共面,所以 $\boldsymbol{a}\times\boldsymbol{b}\perp\boldsymbol{c}$,

$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=0$,即

$$(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=0.$$

反之,当$(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=0$时,由$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=0$可知,$\boldsymbol{c}\perp\boldsymbol{a}\times\boldsymbol{b}$.

若$\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{0}$,则$\boldsymbol{a}/\!/\boldsymbol{b}$.平移后$\boldsymbol{a}$与$\boldsymbol{b}$共线,因而$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$共面.

若$\boldsymbol{a}\times\boldsymbol{b}\neq\boldsymbol{0}$,由$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$都与$\boldsymbol{a}\times\boldsymbol{b}$垂直可知,$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$也共面.

综上所述,该定理的结论成立. □

例 1.5 证明:平行四边形$ABCD$为菱形$\Leftrightarrow$对角线$AC\perp DB$.

证明 设$\overrightarrow{AB}=\boldsymbol{a},\overrightarrow{AD}=\boldsymbol{b}$,于是$\overrightarrow{AC}=\boldsymbol{a}+\boldsymbol{b},\overrightarrow{DB}=\boldsymbol{a}-\boldsymbol{b}$,因为

$$(\boldsymbol{a}+\boldsymbol{b})\cdot(\boldsymbol{a}-\boldsymbol{b})=\boldsymbol{a}\cdot\boldsymbol{a}-\boldsymbol{b}\cdot\boldsymbol{b}=|\boldsymbol{a}|^2-|\boldsymbol{b}|^2,$$

所以,

$$ABCD\text{为菱形}\Leftrightarrow|\boldsymbol{a}|=|\boldsymbol{b}|\Leftrightarrow(\boldsymbol{a}+\boldsymbol{b})\cdot(\boldsymbol{a}-\boldsymbol{b})=0\Leftrightarrow\overrightarrow{AC}\perp\overrightarrow{DB}$$

(即线段$AC\perp DB$). □

例 1.6 当k为何值时,四个点$P(2,0,1),A(1,2,3),B(2,3,1)$和$C(3,1,k)$共面?

解 令$\boldsymbol{a}=\overrightarrow{PA},\boldsymbol{b}=\overrightarrow{PB},\boldsymbol{c}=\overrightarrow{PC}$,则它们的坐标分别为

$$\boldsymbol{a}=(-1,2,2),\quad \boldsymbol{b}=(0,3,0),\quad \boldsymbol{c}=(1,1,k-1),$$

$$(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=\begin{vmatrix}-1&0&1\\2&3&1\\2&0&k-1\end{vmatrix}=-3(k+1).$$

可知,当$k=-1$时,$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$共面,即所给四点共面.

§1.3 空间平面及其方程

本节以向量为工具,在空间直角坐标系中根据确定平面的条件建立相应的平面方程.

1.3.1 平面的点法式方程

与平面垂直的非零向量叫做该**平面的法向量**.显然,平面上任一向量都与其法向量垂直.

由中学几何可知,一个平面由其上一点和它的法向量唯一确定.下面我们在此条件下来建立平面的方程.

设$P_0(x_0,y_0,z_0)$是平面π上一点,$\boldsymbol{n}=A\boldsymbol{i}+B\boldsymbol{j}+C\boldsymbol{k}$为平面$\pi$的法向量.不难得出,$P(x,y,z)$在平面$\pi$上(如图1.8)$\Leftrightarrow\overrightarrow{P_0P}\perp\boldsymbol{n}\Leftrightarrow\overrightarrow{P_0P}\cdot\boldsymbol{n}=0$,即

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0,\tag{1.1}$$

这就是平面 π 的方程,称这个方程为**平面的点法式方程**.

例 1.7 已知平面过点 $P_0(x_0,y_0,z_0)$ 且平行于两个不共线的向量

$$\boldsymbol{a}=a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k} \quad 和 \quad \boldsymbol{b}=b_x\boldsymbol{i}+b_y\boldsymbol{j}+b_z\boldsymbol{k},$$

求此平面的方程.

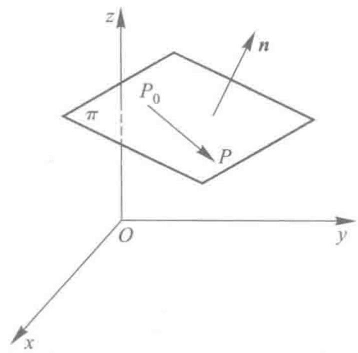


图 1.8

解法 1 因为所求平面平行于向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$,所以平面的法向量同时垂直于 $\boldsymbol{a}$ 和 $\boldsymbol{b}$.从而平面的法向量 $\boldsymbol{n}$ 可取为 $\boldsymbol{a}\times\boldsymbol{b}$,即

$$\boldsymbol{n}=\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix}\boldsymbol{i}-\begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix}\boldsymbol{j}+\begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix}\boldsymbol{k}.$$

由平面的点法式方程得出所求平面方程为

$$\begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix}(x-x_0)-\begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix}(y-y_0)+\begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix}(z-z_0)=0.$$

写成便于记忆的行列式形式,则为

$$\begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}=0.$$

解法 2 设 $P(x,y,z)$ 为该平面上的任一点,则向量 $\overrightarrow{P_0P},\boldsymbol{a},\boldsymbol{b}$ 共面.于是,$(\overrightarrow{P_0P},\boldsymbol{a},\boldsymbol{b})=0$,代入向量 $\overrightarrow{P_0P},\boldsymbol{a},\boldsymbol{b}$ 的坐标,可得该平面的方程为

$$\begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}=0.$$

1.3.2 平面的一般式方程

将平面的点法式方程(1.1)整理,得

$$Ax+By+Cz+D=0, \tag{1.2}$$

其中 $D=-(Ax_0+By_0+Cz_0)$.把这个方程称为**平面的一般式方程**.

可见,任何一个平面都可用一个三元一次方程来表示.反之,任意一个三元一次方程

$$Ax+By+Cz+D=0 \quad (其中\ A,B,C\ 不全为\ 0)$$

总表示一个平面.理由如下:

不失一般性,设 $C\neq0$,则该方程可化为

$$A(x-0)+B(y-0)+C\left[z-\left(-\frac{D}{C}\right)\right]=0,$$

这是一个过点$\left(0,0,-\frac{D}{C}\right)$以 $\boldsymbol{n}=(A,B,C)$为法向量的平面.因此,一个三元一次方程表示一个平面.

从上面分析可知,平面与三元一次方程相对应,且方程中变量的系数组成的向量 $\boldsymbol{n}=(A,B,C)$是平面的法向量.

下面给出特殊的三元一次方程对应的平面的特点:

(1) 当 $D=0$ 时,式(1.2)表示一个过原点的平面.

(2) 当 $D\neq0$ 且 A,B,C 中只有一个为零时,则平面平行于某个坐标轴.

例如,当只有 $A=0$ 时,平面的法向量为 $\boldsymbol{n}=(0,B,C)$.由 $\boldsymbol{n}\cdot\boldsymbol{i}=0$ 知,平面的法向量 $\boldsymbol{n}$ 与 x 轴垂直,因此平面平行于 x 轴.

(3) 当 $D\neq0$ 且 A,B,C 中有两个为零时,则平面平行于某个坐标面.

例如,当 $A=0,B=0$ 时,则平面既平行于 x 轴,又平行于 y 轴,因此平面平行于 Oxy 面.

例 1.8　求过两点 $P(1,0,0)$和 $Q(0,0,1)$且平行于 y 轴的平面方程.

解　由于平面平行于 y 轴,因此平面的方程可设为

$$Ax+Cz+D=0.$$

因为平面过 P 和 Q 两点,将其坐标代入方程,得

$$\begin{cases}A+D=0,\\C+D=0,\end{cases}$$

于是,$A=C=-D$,故所求的平面方程为 $x+z-1=0$.

1.3.3　平面的截距式方程

一个不过原点的平面与 x 轴的交点的横坐标叫做该平面在 x 轴上的截距.类似地,可给出平面在 y 轴和 z 轴上的截距的定义.

当 $abc\neq0$ 时,平面$\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1$ 在 x 轴、y 轴、z 轴上的截距分别为 a,b,c.这种形式的平面方程称为**平面的截距式方程**.

通过截距式方程可以比较容易地画出平面的图形,例如,将平面方程

$$10x+5y+4z-20=0$$

化成截距式方程

$$\frac{x}{2}+\frac{y}{4}+\frac{z}{5}=1.$$

按照三个坐标轴上的截距就可画出平面的图形(图 1.9).

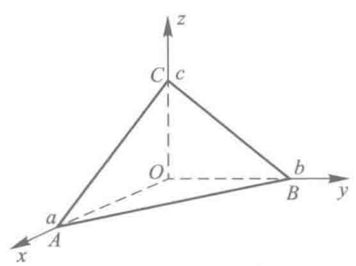


图 1.9

1.3.4 平面的三点式方程

由中学几何可知,空间中不共线的三点可唯一地确定一个平面.下面我们来建立过不共线三点 $P_1(x_1,y_1,z_1)$,$P_2(x_2,y_2,z_2)$,$P_3(x_3,y_3,z_3)$的平面方程.

设 $P(x,y,z)$为平面上的任一点,则向量$\overrightarrow{P_1P}$,$\overrightarrow{P_1P_2}$,$\overrightarrow{P_1P_3}$的坐标分别为$(x-x_1,y-y_1,z-z_1)$,$(x_2-x_1,y_2-y_1,z_2-z_1)$和$(x_3-x_1,y_3-y_1,z_3-z_1)$.

因为这三个向量共面,所以它们的混合积为零.于是

$$\begin{vmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{vmatrix}=0,$$

这就是由三点 P_1,P_2 和 P_3 所确定的平面方程,称为**平面的三点式方程**.

1.3.5 同轴平面束

经过同一条直线的所有平面的集合叫做**同轴平面束**.

当平面 $\pi_1:A_1x+B_1y+C_1z+D_1=0$ 与平面 $\pi_2:A_2x+B_2y+C_2z+D_2=0$ 相交时,它们交于一条直线,记为 l.

设 λ_1 与 λ_2 是两个不同时为零的实参数,建立含有 λ_1 与 λ_2 的三元一次方程

$$\pi_3:\lambda_1(A_1x+B_1y+C_1z+D_1)+\lambda_2(A_2x+B_2y+C_2z+D_2)=0. \tag{1.3}$$

当 $\lambda_1=0$ 时,它是平面 π_2,当 $\lambda_2=0$ 时,它是平面 π_1.

对于给定的 λ_1 和 λ_2,式(1.3)是关于 x,y,z 的三元一次方程,它表示一个平面.由于直线 l 上的任一点的坐标都满足平面 π_1 和 π_2 的方程,当然也满足方程(1.3),因此方程(1.3)表示的平面经过直线 l.另一方面,平面 π_1 的法向量取为 $\boldsymbol{n}_1=(A_1,B_1,C_1)$,平面 π_2 的法向量取为 $\boldsymbol{n}_2=(A_2,B_2,C_2)$,平面 π_3 的法向量取为 $\boldsymbol{n}_3=(\lambda_1A_1+\lambda_2A_2,\lambda_1B_1+\lambda_2B_2,\lambda_1C_1+\lambda_2C_2)$,则 $\boldsymbol{n}_3=\lambda_1\boldsymbol{n}_1+\lambda_2\boldsymbol{n}_2$.再由 λ_1 和 λ_2 为实参数得出,平面 π_3 的法向量可以取遍所有与直线 l 垂直的方向,因此方程(1.3)是以直线 l 为轴的平面束方程.

用平面束方程处理某些问题有时会很方便.

例 1.9 求过点 $P(-1,0,1)$并且经过两个平面 $x+3y-z=0$ 和 $x-y+z+1=0$ 的交线的平面的方程.

解 由于所求平面与两个已知平面均为同轴平面束中的平面,因此可设所求平面为

$$\lambda_1(x+3y-z)+\lambda_2(x-y+z+1)=0,$$

将 $P(-1,0,1)$代入方程,得出 $\lambda_2=2\lambda_1$.将 $\lambda_2=2\lambda_1$ 代入上式并消去 λ_1,得到所求的平面方程为

$$3x+y+z+2=0.$$

§ 1.4 空间直线及其方程

本节仍以向量为工具,在空间直角坐标系中根据确定直线的条件建立相应的直线方程.

1.4.1 直线的点向式方程与参数式方程

与直线平行的非零向量叫做该**直线的方向向量**,它的三个坐标称为该**直线的方向数**.显然,直线上的任一个向量都与其方向向量平行.

由中学几何可知,一条直线由其上一点和它的方向向量唯一确定.下面我们在此条件下来建立直线的方程.

设 $P_0(x_0,y_0,z_0)$ 是直线 l 上一点,$\boldsymbol{s}=m\boldsymbol{i}+n\boldsymbol{j}+p\boldsymbol{k}$ 为 l 的方向向量,即方向数为 m,n,p.

点 $P(x,y,z)$ 在直线 l 上 $\Leftrightarrow \overrightarrow{P_0P} /\!/ \boldsymbol{s}$,即

$$\frac{x-x_0}{m}=\frac{y-y_0}{n}=\frac{z-z_0}{p}, \tag{1.4}$$

这种由一点和方向向量所确定的直线方程叫做**直线的点向式方程**.由于方程的结构对称,因此也叫做**直线的对称式方程**.

因为式(1.4)的分母表示的是直线的方向向量的坐标,所以当方向数 m,n,p 中有的为零时,在形式上仍可用式(1.4)来表示直线.例如

$$\frac{x-1}{2}=\frac{y-2}{0}=\frac{z-3}{-3}$$

表示一条过点(1,2,3),且以(2,0,-3)为方向向量的直线.

对于上面的直线 l,我们现在来建立它的参数式方程.

点 $P(x,y,z)$ 在直线 l 上 $\Leftrightarrow \overrightarrow{P_0P} /\!/ \boldsymbol{s} \Leftrightarrow$ 存在实数 t,使得 $\overrightarrow{P_0P}=t\boldsymbol{s}$,即

$$(x-x_0,y-y_0,z-z_0)=t(m,n,p),$$

也即

$$\begin{cases} x=x_0+mt, \\ y=y_0+nt, \\ z=z_0+pt, \end{cases} \tag{1.5}$$

这个方程叫做直线的参数式方程,其中 t 为参数.

注意 令式(1.4)等于参数 t,可将直线的对称式方程化为参数式方程;反过来,从式(1.5)的三个式子分别求出 t,可将直线的参数式方程化为对称式方程.

例 1.10 求过两个相异点 $P_0(x_0,y_0,z_0)$ 和 $P_1(x_1,y_1,z_1)$ 的直线方程.

解 所求直线是过点 P_0 且以 $\overrightarrow{P_0P_1}$ 为方向向量的直线.

由 $\overrightarrow{P_0P_1}$ 的坐标向量 $(x_1-x_0,y_1-y_0,z_1-z_0)$ 可直接写出该直线的方程

$$\frac{x-x_0}{x_1-x_0}=\frac{y-y_0}{y_1-y_0}=\frac{z-z_0}{z_1-z_0},$$

这种由两点确定的直线方程叫做直线的两点式方程.

例 1.11 求直线 $\frac{x-1}{2}=\frac{y-2}{0}=\frac{z-3}{-3}$ 与平面 $2x-y+z=4$ 的交点 P 的坐标.

解 交点 P 的三个坐标是直线方程和平面方程的公共解.将直线写成参数式方程

$$\begin{cases}x=1+2t,\\y=2,\\z=3-3t,\end{cases}$$

代入平面方程 $2x-y+z=4$,求得 $t=1$.因此所求的交点为 $P(3,2,0)$.

1.4.2 直线的一般式方程

当平面 $\pi_1:A_1x+B_1y+C_1z+D_1=0$ 与平面 $\pi_2:A_2x+B_2y+C_2z+D_2=0$ 相交时,它们的交线 l 上的点的坐标同时满足 π_1 与 π_2 的方程,即满足方程组

$$\begin{cases}A_1x+B_1y+C_1z+D_1=0,\\A_2x+B_2y+C_2z+D_2=0.\end{cases}$$

反过来,坐标满足该方程组的点同时在平面 π_1 与 π_2 上,即在交线 l 上.因此,可用该方程组表示直线 l,称为**直线 l 的一般式方程**.

给定两个相交的平面,它们所交的直线就唯一确定.反之,给定一条直线,由于过该直线的平面有无穷多个,将其中任何两个方程联立得到的方程组都可表示这条直线.因此,直线的一般式方程不唯一.

上述几种形式的直线方程,对于处理不同的问题各有其长处,应熟悉它们相互间的转化.

对于方程(1.4),当 m,n,p 中有的数为零时,将对称式方程转化为一般式方程的做法如下:

当 m,n,p 中只有一个为零,例如 $m=0$ 时,式(1.4)可化成一般式方程

$$\begin{cases}x=x_0,\\\frac{y-y_0}{n}=\frac{z-z_0}{p}.\end{cases}$$

当 m,n,p 中有两个为零,例如 $n=0,p=0$ 时,式(1.4)可化成一般式方程

$$\begin{cases} y = y_0, \\ z = z_0. \end{cases}$$

当 m,n,p 都不为零时，将直线的对称式方程化为一般式方程的做法请大家自己考虑. 下面我们来介绍将直线的一般式方程化为对称式方程的做法.

例 1.12 将直线的一般式方程$\begin{cases} x-y+z=0, \\ x+y-2z=2 \end{cases}$化为对称式方程.

解法 1 先求直线上一点. 在该直线方程中，令 $z=0$，得

$$\begin{cases} x - y = 0, \\ x + y = 2, \end{cases}$$

求得 $x=y=1$. 因此，点(1,1,0)是该直线上的一个点.

再求该直线的方向向量 $\boldsymbol{s}$. 由于这条直线是法向量分别为

$$\boldsymbol{n}_1 = (1,-1,1) \quad 和 \quad \boldsymbol{n}_2 = (1,1,-2)$$

的两个平面的交线，因此 $\boldsymbol{s} \perp \boldsymbol{n}_1$ 且 $\boldsymbol{s} \perp \boldsymbol{n}_2$. 取

$$\boldsymbol{s} = \boldsymbol{n}_1 \times \boldsymbol{n}_2 = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 1 & -1 & 1 \\ 1 & 1 & -2 \end{vmatrix} = \boldsymbol{i} + 3\boldsymbol{j} + 2\boldsymbol{k},$$

于是，该直线的对称式方程为

$$\frac{x-1}{1} = \frac{y-1}{3} = \frac{z}{2}.$$

解法 2 令 $z=0$，求得该直线上的一点(1,1,0)；再令 $y=0$，又求得该直线上的一点 $\left(\frac{2}{3},0,-\frac{2}{3}\right)$.

根据两点式方程可得该直线的对称式方程为

$$\frac{x-1}{-\frac{1}{3}} = \frac{y-1}{-1} = \frac{z}{-\frac{2}{3}},$$

即$\frac{x-1}{1} = \frac{y-1}{3} = \frac{z}{2}$.

解法 3 在$\begin{cases} x-y+z=0, \\ x+y-2z=2 \end{cases}$中，令 $z=t$，求出 x 和 y，可得该直线的参数式方程

$$\begin{cases} x = \frac{1}{2}t + 1, \\ y = \frac{3}{2}t + 1, \\ z = t, \end{cases}$$

再将其化为对称式方程,得$\frac{x-1}{1}=\frac{y-1}{3}=\frac{z}{2}$.

§1.5 位置关系、夹角与距离

本节研究平面与平面、直线与平面及直线与直线间的关系,并介绍各种夹角和距离的计算方法.

1.5.1 两平面间的关系

两个平面之间有重合、平行和相交这三种关系,下面给出其判别方法.设有两个平面

$$\pi_1: A_1x + B_1y + C_1z + D_1 = 0,$$

$$\pi_2: A_2x + B_2y + C_2z + D_2 = 0,$$

它们的法向量分别为

$$\boldsymbol{n}_1 = (A_1, B_1, C_1) \text{ 和 } \boldsymbol{n}_2 = (A_2, B_2, C_2).$$

当 $\boldsymbol{n}_1 /\!/ \boldsymbol{n}_2$,即$\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}$时,平面 π_1 与 π_2 或者重合,或者平行.因此,有下面结论:

(1) 平面 π_1 与 π_2 重合$\Leftrightarrow\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}=\frac{D_1}{D_2}$;

(2) 平面 π_1 与 π_2 平行$\Leftrightarrow\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}\neq\frac{D_1}{D_2}$;

(3) 平面 π_1 与 π_2 相交$\Leftrightarrow A_1 : B_1 : C_1 \neq A_2 : B_2 : C_2$$\left(\text{即}\frac{A_1}{A_2},\frac{B_1}{B_2},\frac{C_1}{C_2}\text{不全相等}\right)$.

1.5.2 直线与平面间的关系

下面讨论一条直线 l 与一个平面 π 之间的关系,设直线 l 与平面 π 的方程分别为

$$l: \frac{x-x_0}{m}=\frac{y-y_0}{n}=\frac{z-z_0}{p},$$

$$\pi: Ax + By + Cz + D = 0,$$

它们的方向向量和法向量分别为

$$\boldsymbol{s} = (m, n, p)$$

和

$$n=(A,B,C),$$

(1) l 在 π 上$\Leftrightarrow s\perp n$(即 $s\cdot n=0$)且点(x_0,y_0,z_0)满足 π 的方程$\Leftrightarrow mA+nB+pC=0$ 且 $Ax_0+By_0+Cz_0+D=0$.

(2) l 与 π 平行$\Leftrightarrow s\perp n$ 且点(x_0,y_0,z_0)不满足 π 的方程$\Leftrightarrow mA+nB+pC=0$ 且 $Ax_0+By_0+Cz_0+D\neq 0$.

(3) l 与 π 相交$\Leftrightarrow s$ 与 n 不垂直$\Leftrightarrow mA+nB+pC\neq 0$.

(4) l 与 π 垂直$\Leftrightarrow s/\!/n\Leftrightarrow\frac{A}{m}=\frac{B}{n}=\frac{C}{p}$.

1.5.3 两直线间的关系

设有两条空间直线

$$l_1:\frac{x-x_1}{m_1}=\frac{y-y_1}{n_1}=\frac{z-z_1}{p_1},$$

$$l_2:\frac{x-x_2}{m_2}=\frac{y-y_2}{n_2}=\frac{z-z_2}{p_2}.$$

l_1 过点 $P_1(x_1,y_1,z_1)$且方向向量为 $s_1=(m_1,n_1,p_1)$, l_2 过点 $P_2(x_2,y_2,z_2)$且方向向量为 $s_2=(m_2,n_2,p_2)$,向量$\overrightarrow{P_1P_2}$的坐标向量为

$$p=(x_2-x_1,y_2-y_1,z_2-z_1).$$

我们首先研究直线 l_1 与 l_2 共面的判别方法.

直线 l_1 与 l_2 共面$\Leftrightarrow$向量 $s_1,s_2,\overrightarrow{P_1P_2}$共面$\Leftrightarrow(s_1,s_2,\overrightarrow{P_1P_2})=0\Leftrightarrow(s_1,s_2,p)=0$,即

$$\begin{vmatrix} m_1 & m_2 & x_2-x_1 \\ n_1 & n_2 & y_2-y_1 \\ p_1 & p_2 & z_2-z_1 \end{vmatrix}=0. \tag{1.6}$$

当 l_1 与 l_2 共面时,它们之间有重合、平行和相交三种情况.其判别如下:

(1) l_1 与 l_2 重合$\Leftrightarrow s_1/\!/s_2/\!/\overrightarrow{P_1P_2}\Leftrightarrow m_1:n_1:p_1=m_2:n_2:p_2=(x_2-x_1):(y_2-y_1):(z_2-z_1)$.

(2) l_1 与 l_2 平行$\Leftrightarrow s_1/\!/s_2$ 且不平行于$\overrightarrow{P_1P_2}$

$$\Leftrightarrow m_1:n_1:p_1=m_2:n_2:p_2\neq(x_2-x_1):(y_2-y_1):(z_2-z_1).$$

(3) l_1 与 l_2 交于一点$\Leftrightarrow s_1,s_2,\overrightarrow{P_1P_2}$共面且 s_1 与 s_2 不平行$\Leftrightarrow$式(1.6)成立并且 $m_1:n_1:p_1\neq m_2:n_2:p_2$.

若式(1.6)不成立,则直线 l_1 与 l_2 异面.

例 1.13 设两条直线分别为

$$l_1:\frac{x-1}{-1}=\frac{y+2}{2}=\frac{z-1}{1},$$

$$l_2:\frac{x-r}{2}=\frac{y-2}{t}=\frac{z+1}{-2}.$$

对 r 和 t 加以讨论,分析 l_1 与 l_2 在空间中的位置关系.

解 l_1 过点 $P_1(1,-2,1)$ 且方向向量为 $\boldsymbol{s}_1=(-1,2,1)$,$l_2$ 过点 $P_2(r,2,-1)$ 且方向向量为 $\boldsymbol{s}_2=(2,t,-2)$,$\overrightarrow{P_1P_2}$的坐标向量为 $\boldsymbol{p}=(r-1,4,-2)$,由

$$|\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{p}|=\begin{vmatrix}-1 & 2 & r-1\\ 2 & t & 4\\ 1 & -2 & -2\end{vmatrix}=-(r-3)(t+4),$$

得出

(1) 当 $r\neq3$ 且 $t\neq-4$ 时,$|\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{p}|\neq0$,$l_1$ 与 l_2 异面.

(2) 当 $t=-4$ 时,对 $\forall r\in\mathbb{R}$,都有

$(-1):2:1=2:(-4):(-2)\neq(r-1):4:(-2)$,因此 l_1 与 l_2 平行.

(3) 当 $t\neq-4$ 且 $r=3$ 时,由 $|\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{p}|=0$ 且 $(-1):2:1\neq2:t:(-2)$,得出 l_1 与 l_2 相交.

1.5.4 直线和平面相互间的夹角

现在我们通过向量的夹角来求直线与直线、平面与平面以及直线与平面的夹角.

设直线 l_1 和 l_2 的方向向量分别为 $\boldsymbol{s}_1$ 和 $\boldsymbol{s}_2$,平面 π_1 和 π_2 的法向量分别为 $\boldsymbol{n}_1$ 和 $\boldsymbol{n}_2$.

(1) 设 $\boldsymbol{s}_1$ 与 $\boldsymbol{s}_2$ 的夹角为 θ,则直线 l_1 与 l_2 的夹角为 $\varphi=\min\{\theta,\pi-\theta\}$,

$$\cos\varphi=|\cos\theta|=\frac{|\boldsymbol{s}_1\cdot\boldsymbol{s}_2|}{|\boldsymbol{s}_1||\boldsymbol{s}_2|}.$$

(2) 设 $\boldsymbol{n}_1$ 和 $\boldsymbol{n}_2$ 的夹角为 θ,则平面 π_1 和 π_2 的夹角为 $\varphi=\min\{\theta,\pi-\theta\}$,

$$\cos\varphi=|\cos\theta|=\frac{|\boldsymbol{n}_1\cdot\boldsymbol{n}_2|}{|\boldsymbol{n}_1||\boldsymbol{n}_2|}.$$

(3) 设 $\boldsymbol{s}_1$ 和 $\boldsymbol{n}_1$ 的夹角为 θ,则直线 l_1 和平面 π_1 的夹角为 $\varphi=\frac{\pi}{2}-\min\{\theta,\pi-\theta\}$(如图 1.10),

$$\sin\varphi=\cos(\min\{\theta,\pi-\theta\})=|\cos\theta|=\frac{|\boldsymbol{s}_1\cdot\boldsymbol{n}_1|}{|\boldsymbol{s}_1||\boldsymbol{n}_1|}.$$

注意 上述三种夹角都不大于$\frac{\pi}{2}$,这与向量的夹角有所不同.

例 1.14 求直线 $l:\frac{x-1}{-1}=\frac{y+2}{2}=\frac{z-1}{1}$和平面 $\pi:x+y+2z=3$ 的夹角 φ.

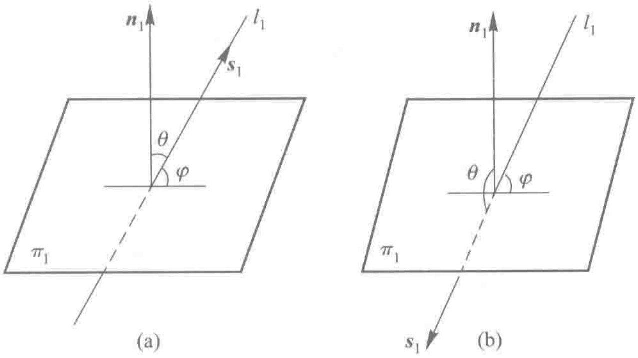


图 1.10

解 直线 l 的方向向量为 $\boldsymbol{s}=(-1,2,1)$，平面 π 的法向量为 $\boldsymbol{n}=(1,1,2)$. 由

$$\boldsymbol{s}\cdot\boldsymbol{n}=(-1)\times 1+2\times 1+1\times 2=3,$$

$$|\boldsymbol{s}|=\sqrt{(-1)^2+2^2+1^2}=\sqrt{6},$$

$$|\boldsymbol{n}|=\sqrt{1^2+1^2+2^2}=\sqrt{6},$$

得

$$\sin\varphi=\frac{|\boldsymbol{s}\cdot\boldsymbol{n}|}{|\boldsymbol{s}||\boldsymbol{n}|}=\frac{1}{2},\quad \text{因此 } \varphi=\frac{\pi}{6}.$$

1.5.5 距离

下面我们仍以向量为工具研究点到直线、点到平面以及两条异面直线之间的距离.由前两种距离还可求出两条平行直线、两个平行平面以及直线与平行于该直线的平面之间的距离.

1. 点到直线的距离

设 $P_0(x_0,y_0,z_0)$ 为空间一点，l 为过点 $P_1(x_1,y_1,z_1)$ 且方向向量为 $\boldsymbol{s}$ 的直线（如图 1.11 所示），点 P_0 到直线 l 的距离用 $d(P_0,l)$ 表示.

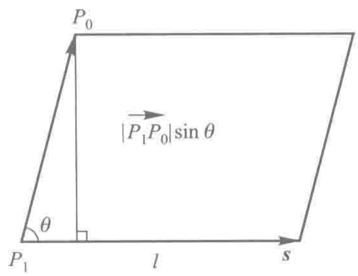


图 1.11

显然，$d(P_0,l)=|P_1P_0|\sin\theta$（其中 θ 为 $\boldsymbol{s}$ 与 $\overrightarrow{P_1P_0}$ 的夹角），由

$$|\boldsymbol{s}\times\overrightarrow{P_1P_0}|=|\boldsymbol{s}||\overrightarrow{P_1P_0}|\sin\theta,$$

可得

$$d(P_0,l)=\frac{|\boldsymbol{s}\times\overrightarrow{P_1P_0}|}{|\boldsymbol{s}|}.$$

不难得出，两条平行直线的距离就是其中一条直线上的一点到另一条直线的距离.

例 1.15 求两条平行直线 l_1 和 l_2 的距离 $d(l_1,l_2)$,其中

$$l_1:\frac{x}{1}=\frac{y-2}{2}=\frac{z-1}{2},$$

$$l_2:\frac{x-1}{2}=\frac{y-2}{4}=\frac{z-2}{4}.$$

解 直线 l_1 过点 $P_1(0,2,1)$,方向向量记为 $\boldsymbol{s}_1$,直线 l_2 过点 $P_2(1,2,2)$.由 $\overrightarrow{P_1P_2}=\boldsymbol{i}+\boldsymbol{k}$,得

$$\boldsymbol{s}_1\times\overrightarrow{P_1P_2}=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ 1 & 2 & 2\\ 1 & 0 & 1\end{vmatrix}=2\boldsymbol{i}+\boldsymbol{j}-2\boldsymbol{k},$$

再由 $|\boldsymbol{s}_1\times\overrightarrow{P_1P_2}|=3$ 及 $|\boldsymbol{s}_1|=3$,得

$$d(P_2,l_1)=\frac{|\boldsymbol{s}_1\times\overrightarrow{P_1P_2}|}{|\boldsymbol{s}_1|}=\frac{3}{3}=1,$$

因此直线 l_1 和 l_2 的距离为 1.

2. 点到平面的距离

设平面 π 的方程为 $Ax+By+Cz+D=0$,$P_1(x_1,y_1,z_1)$ 为平面 π 上的一点,则平面 π 的法向量为 $\boldsymbol{n}=(A,B,C)$.另设 $P_0(x_0,y_0,z_0)$ 为平面 π 外的一点,过 P_0 作平面 π 的垂线与 π 交于点 Q_0.(图 1.12).

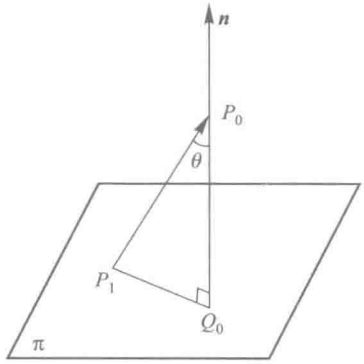


图 1.12

用 $d(P_0,\pi)$ 表示点 P_0 到平面 π 的距离,则

$$d(P_0,\pi)=|\overrightarrow{Q_0P_0}|=|\overrightarrow{P_1P_0}||\cos\theta|\quad(\text{其中 }\theta\text{ 为 }\boldsymbol{n}\text{ 与 }\overrightarrow{P_1P_0}\text{ 的夹角}).$$

由 $|\boldsymbol{n}\cdot\overrightarrow{P_1P_0}|=|\boldsymbol{n}||\overrightarrow{P_1P_0}||\cos\theta|$,得

$$d(P_0,\pi)=\frac{|\boldsymbol{n}\cdot\overrightarrow{P_1P_0}|}{|\boldsymbol{n}|}=\frac{|A(x_0-x_1)+B(y_0-y_1)+C(z_0-z_1)|}{\sqrt{A^2+B^2+C^2}}.$$

由 P_1 在平面 π 上可知,

$$Ax_1+By_1+Cz_1+D=0,$$

即

$$-Ax_1-By_1-Cz_1=D.$$

因此,点 P_0 到平面 π 的距离为

$$d(P_0,\pi)=\frac{|Ax_0+By_0+Cz_0+D|}{\sqrt{A^2+B^2+C^2}}.$$

不难得出:

(1) 与平面 π 平行的直线 l 到 π 的距离为 l 上的任一点到 π 的距离.

(2) 两个平行平面的距离为其中一个平面上的任一点到另一个平面的距离.

例 1.16　求两个平行平面 π_1 和 π_2 之间的距离 $d(\pi_1,\pi_2)$，其中

$$\pi_1: x+2y-2z+3=0,$$
$$\pi_2: 2x+4y-4z-3=0.$$

解　在 π_1 上取一点 $P(-3,0,0)$，则

$$d(\pi_1,\pi_2)=\frac{|2\times(-3)+4\times 0-4\times 0-3|}{\sqrt{2^2+4^2+(-4)^2}}=\frac{3}{2}.$$

***3. 两条异面直线的距离**

设直线 l_1 和 l_2 为异面直线，$\boldsymbol{s}_1$ 与 $\boldsymbol{s}_2$ 分别为它们的方向向量. 记 $\boldsymbol{n}=\boldsymbol{s}_1\times\boldsymbol{s}_2$，用 $d(l_1,l_2)$ 表示直线 l_1 和 l_2 的距离.

设 P_1 为直线 l_1 上的一点，过点 P_1 做一个法向量为 $\boldsymbol{n}$ 的平面 π，显然 π 与 l_2 平行. 于是，直线 l_1 和 l_2 的距离就是直线 l_2 到平面 π 的距离，亦即 l_2 上的一点 P_2 到 π 的距离. 因此

$$d(l_1,l_2)=d(l_2,\pi)=d(P_2,\pi)$$
$$=\frac{|\boldsymbol{n}\cdot\overrightarrow{P_1P_2}|}{|\boldsymbol{n}|}=\frac{|(\boldsymbol{s}_1\times\boldsymbol{s}_2)\cdot\overrightarrow{P_1P_2}|}{|\boldsymbol{s}_1\times\boldsymbol{s}_2|}=\frac{|(\boldsymbol{s}_1,\boldsymbol{s}_2,\overrightarrow{P_1P_2})|}{|\boldsymbol{s}_1\times\boldsymbol{s}_2|}.$$

因为 $|(\boldsymbol{s}_1,\boldsymbol{s}_2,\overrightarrow{P_1P_2})|$ 表示以 $\boldsymbol{s}_1,\boldsymbol{s}_2,\overrightarrow{P_1P_2}$ 为相邻棱的平行六面体的体积，$|\boldsymbol{s}_1\times\boldsymbol{s}_2|$ 表示该平行六面体的底面积，所以异面直线 l_1 和 l_2 的距离等于该平行六面体的高.

与异面直线 l_1 和 l_2 都垂直相交的直线称为异面直线的公垂线，它是过 l_1 且平行于向量 $\boldsymbol{n}=\boldsymbol{s}_1\times\boldsymbol{s}_2$ 的平面与过 l_2 也平行于向量 $\boldsymbol{n}$ 的平面的交线，这个交线在 l_1 与 l_2 之间的线段长度就是直线 l_1 和 l_2 之间的距离 $d(l_1,l_2)$.

习　题　一

1. 判断下列等式何时成立：

(1) $|\boldsymbol{a}+\boldsymbol{b}|=|\boldsymbol{a}-\boldsymbol{b}|$；

(2) $|\boldsymbol{a}+\boldsymbol{b}|=|\boldsymbol{a}|+|\boldsymbol{b}|$；

(3) $|\boldsymbol{a}+\boldsymbol{b}|=|\boldsymbol{a}|-|\boldsymbol{b}|$；

(4) $\dfrac{\boldsymbol{a}}{|\boldsymbol{a}|}=\dfrac{\boldsymbol{b}}{|\boldsymbol{b}|}$.

2. 设 $\boldsymbol{a}\neq\boldsymbol{0}$，下列结论成立吗？

(1) 若 $\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{a}\cdot\boldsymbol{c}$，则 $\boldsymbol{b}=\boldsymbol{c}$；

(2) 若 $\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{a}\times\boldsymbol{c}$，则 $\boldsymbol{b}=\boldsymbol{c}$；

(3) 若 $\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{a}\cdot\boldsymbol{c}$ 且 $\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{a}\times\boldsymbol{c}$，则 $\boldsymbol{b}=\boldsymbol{c}$；

(4) $(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}=\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c})$；

(5) $(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}=\boldsymbol{c}\times(\boldsymbol{a}\times\boldsymbol{b})$;

(6) $(\boldsymbol{a}\times\boldsymbol{b})\times\boldsymbol{c}$ 与 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 共面.

3. 已知三点 $A(x_1,y_1,z_1)$, $B(x_2,y_2,z_2)$, $C(x_3,y_3,z_3)$, 求 $\triangle ABC$ 的面积.

4. 已知四点 $A_i(x_i,y_i,z_i)(i=1,2,3,4)$, 求四面体 $A_1A_2A_3A_4$ 的体积.

5. 在空间直角坐标系下,

(1) $\boldsymbol{a}=3\boldsymbol{i}-6\boldsymbol{k}$, $\boldsymbol{b}=2\boldsymbol{i}-4\boldsymbol{j}$, 求 $\boldsymbol{a}\cdot\boldsymbol{b}$ 及 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角;

(2) $\boldsymbol{a}=5\boldsymbol{i}+2\boldsymbol{j}+5\boldsymbol{k}$, $\boldsymbol{b}=2\boldsymbol{i}-\boldsymbol{j}+2\boldsymbol{k}$, 求 $\boldsymbol{a}$ 在 $\boldsymbol{b}$ 上的投影及 $\boldsymbol{b}$ 的方向余弦.

6. 若 $\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}=\boldsymbol{0}$, 证明 $\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{b}\times\boldsymbol{c}=\boldsymbol{c}\times\boldsymbol{a}$.

7. 若 $\boldsymbol{a}\times\boldsymbol{b}+\boldsymbol{b}\times\boldsymbol{c}+\boldsymbol{c}\times\boldsymbol{a}=\boldsymbol{0}$, 证明 $\boldsymbol{a}$, $\boldsymbol{b}$ 和 $\boldsymbol{c}$ 共面.

8. 求过点 $(1,0,-1)$ 且平行于向量 $\boldsymbol{a}=2\boldsymbol{i}+\boldsymbol{j}+\boldsymbol{k}$ 和 $\boldsymbol{b}=\boldsymbol{i}-\boldsymbol{j}$ 的平面方程.

9. 求过点 $(1,1,1)$ 且垂直于平面 $x-y+z=7$ 和 $3x+2y-12z+5=0$ 的平面方程.

10. 求过两点 $(4,0,-2)$ 和 $(5,1,7)$ 且平行于 x 轴的平面方程.

11. 求过直线 $\begin{cases}x=2+3t,\\ y=2+t,\\ z=1+2t\end{cases}$ 和点 $(1,2,-1)$ 的平面方程.

12. 求下列直线的参数式方程和对称式方程:

(1) 过点 $P(2,-1,3)$ 且平行于向量 $\boldsymbol{v}=(3,-1,4)$;

(2) 过点 $P(2,3,-5)$ 且与直线 $\dfrac{x-2}{-1}=\dfrac{y}{3}=\dfrac{z+1}{4}$ 平行;

(3) 过两点 $A(-4,-1,3)$ 和 $B(1,3,-2)$;

(4) 过点 $(0,1,2)$, 平行于平面 $x+y+z=2$ 且垂直于直线 $\begin{cases}x=1+t,\\ y=1-t,\\ z=2t.\end{cases}$

13. 求点 $M(1,2,-5)$ 到直线 $l:\dfrac{x-7}{3}=\dfrac{y-2}{2}=\dfrac{z-1}{-2}$ 的距离及垂线方程.

14. 判断直线与平面的位置关系, 若有交点, 求出交点坐标.

(1) $\dfrac{x-5}{2}=\dfrac{y+3}{-2}=\dfrac{z-1}{3}$ 和 $x+2y-5z-11=0$;

(2) $\dfrac{x-3}{2}=\dfrac{y+1}{-5}=\dfrac{z}{3}$ 和 $2x-y-2z+1=0$.

15. 证明直线 $l_1:\dfrac{x-7}{3}=\dfrac{y-2}{2}=\dfrac{z-1}{-2}$ 和 $l_2:\begin{cases}x=1+2t,\\ y=-2-3t,\\ z=5+4t\end{cases}$ 共面, 并求它们所在平面的方程.

16. 求两平面 $x+y+2z=3$ 和 $2x-y+z+2=0$ 的夹角.

17. 求直线$\begin{cases}x+y+3z=0,\\x-y-z=0\end{cases}$与平面 $x-y-z+1=0$ 的夹角.

18. 求直线$\dfrac{x-1}{1}=\dfrac{y}{-1}=\dfrac{z-1}{1}$与直线$\begin{cases}x=1+t,\\y=2+t,\\z=3+2t\end{cases}$的夹角.

第二章 线性方程组

§2.1 线性方程组的解法

前一章中为了考察平面间的位置关系需要解三元一次方程组.本章就来讨论一般的多元一次方程组.

一次方程组又称为线性方程组.设 F 是一个数域,F 上有 n 个未知量 $x_1,x_2,\cdots,x_n$ 的线性方程组形如

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m, \end{cases} \tag{2.1}$$

其中 $a_{ij},b_i\in F(1\leqslant i\leqslant m,1\leqslant j\leqslant n)$.$a_{ij}(1\leqslant i\leqslant m,1\leqslant j\leqslant n)$称为系数,$b_i(1\leqslant i\leqslant m)$称为常数项,或者说 a_{ij}是第 i 个方程中未知量 x_j 的系数,b_i 为第 i 个方程的常数项.方程个数为 m,它与未知量个数 n 没有关系.

方程组(2.1)的一个解即一个有序数组$(c_1,c_2,\cdots,c_n)$,其中 $c_i\in F,i=1,2,\cdots,n$,使得把(2.1)中的未知量代入 $x_1=c_1,x_2=c_2,\cdots,x_n=c_n$ 时,每个方程都成为恒等式.

研究线性方程组,自然我们关心的是如何求解线性方程组.理论上的问题有:如何判断一个给定的线性方程组是否有解,有解时解的个数是多少,解之间的关系如何等等.下面先来讨论如何求解线性方程组.

初中时大家已解过二元一次方程组或三元一次方程组.所用方法为消元法,包括加减消元、代入消元等.解一般的线性方程组的方法也是如此.回忆一下,当时为了消元,通常是用一个非零数乘以某一个方程或者把一个方程的倍数加到另一个方程上去.我们把这些熟悉的方程组中方程的变形总结成下面三种线性方程组的初等变换,其中出现的数都属于数域 F:

(1) 交换两个方程的位置.

(2) 用一个非零数乘以一个方程.

(3) 把一个方程的倍数加到另一个方程上.

先证明线性方程组的初等变换不改变方程组的解.只以第 3 种初等变换为例.设把方程组(2.1)的第 i 个方程的 k 倍加到第 j 个方程上得到方程组(2.2),则(2.2)与(2.1)只有第 j 个方程不同.方程组(2.2)的第 j 个方程为

$$(ka_{i1}+a_{j1})x_1+(ka_{i2}+a_{j2})x_2+\cdots+(ka_{in}+a_{jn})x_n=kb_i+b_j. \tag{2.3}$$

设$(c_1,c_2,\cdots,c_n)$是方程组(2.1)的一个解,则有

$$a_{i1}c_1+a_{i2}c_2+\cdots+a_{in}c_n=b_i,$$

$$a_{j1}c_1+a_{j2}c_2+\cdots+a_{jn}c_n=b_j,$$

所以

$$(ka_{i1}+a_{j1})c_1+(ka_{i2}+a_{j2})c_2+\cdots+(ka_{in}+a_{jn})c_n=kb_i+b_j.$$

这表明$(c_1,c_2,\cdots,c_n)$也满足方程组(2.2)的第 j 个方程.而(2.2)的其余方程与(2.1)一样.所以$(c_1,c_2,\cdots,c_n)$是方程组(2.2)的一个解.反之,由于方程组的初等变换是可逆的.例如,把方程组(2.2)的第 i 个方程的 $-k$ 倍加到第 j 个方程上就得到方程组(2.1).由前面的证明知,方程组(2.2)的每个解,也是方程组(2.1)的解.对于线性方程组的第 1、2 种初等变换类似地可以证明.

上面证明了若对一个线性方程组做初等变换得到一个新的方程组,则这两个线性方程组同解.

显然可设方程组(2.1)中 x_1 的系数不全为 0,由于可以交换方程的位置,故不妨设 $a_{11}\neq 0$,对 $2\leqslant j\leqslant m$,把第 1 个方程的 $-\dfrac{a_{j1}}{a_{11}}$ 倍加到第 j 个方程上可把第 j 个方程的 x_1 的系数变成 0,即消去未知量 x_1.按这个步骤一直做下去,便可求解线性方程组.

下面看一个例子.用 ⓘ↔ⓙ 表示交换 i,j 这两个方程.用 c · ⓘ 表示用非零数 c 乘以第 i 个方程.用 k · ⓘ+ⓙ 表示把第 i 个方程的 k 倍加到第 j 个方程上.

例 2.1　解线性方程组

$$\begin{cases} 3x_2+x_3=9, \\ x_1-2x_2+x_3=0, \\ 3x_1+3x_2-x_3=6. \end{cases}$$

解

$$\begin{cases} 3x_2+x_3=9, \\ x_1-2x_2+x_3=0, \\ 3x_1+3x_2-x_3=6. \end{cases} \xrightarrow{①\leftrightarrow②} \begin{cases} x_1-2x_2+x_3=0, \\ 3x_2+x_3=9, \\ 3x_1+3x_2-x_3=6. \end{cases}$$

$$\xrightarrow{-3\cdot①+③}\begin{cases}x_1-2x_2+x_3=0,\\ \quad 3x_2+x_3=9,\\ \quad 9x_2-4x_3=6.\end{cases}\xrightarrow{-3\cdot②+③}\begin{cases}x_1-2x_2+x_3=0,\\ \quad 3x_2+x_3=9,\\ \quad -7x_3=-21.\end{cases}$$

$$\xrightarrow{-\frac{1}{7}\cdot③}\begin{cases}x_1-2x_2+x_3=0,\\ \quad 3x_2+x_3=9,\\ \quad x_3=3.\end{cases}\xrightarrow[(-1)\cdot③+②]{(-1)\cdot③+①}\begin{cases}x_1-2x_2=-3,\\ \quad 3x_2=6,\\ \quad x_3=3.\end{cases}$$

$$\xrightarrow{\frac{1}{3}\cdot②}\begin{cases}x_1-2x_2=-3,\\ \quad x_2=2,\\ \quad x_3=3.\end{cases}\xrightarrow{2\cdot②+①}\begin{cases}x_1=1,\\ x_2=2,\\ x_3=3.\end{cases}$$

由此得到原方程组的解为(1,2,3),或写成$\begin{cases}x_1=1,\\ x_2=2,\\ x_3=3.\end{cases}$

在上面例子中可以看到,在解方程组时,只有方程组的系数和常数项参与了运算,未知量在其中只标记了系数所在的位置,所以为简化写法,可以只把方程组的系数和常数项写出,排成矩阵的形式,由此下面先给出矩阵的概念.

定义 2.1 给定数域 F 中 mn 个数 a_{ij},$1\leqslant i\leqslant m$,$1\leqslant j\leqslant n$.将它们排成矩形表,两边用圆括号(或方括号)包起来,即形如

$$\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix},$$

称其为**矩阵**(或数域 F 上的矩阵),通常用大写黑体英文字母表示矩阵.其中的横排称为行,竖排称为列.有m 行 n 列的矩阵称为 $m\times n$ 矩阵.这 mn 个数 a_{ij} 称为矩阵的元素,上述矩阵也可记为$(a_{ij})_{m\times n}$.每个元素都为 0 的矩阵称为**零矩阵**,记为 $\boldsymbol{O}$.数域 F 上所有 $m\times n$ 矩阵构成的集合记为 $M_{m,n}(F)$.n 行 n 列矩阵称为 n 阶方阵.数域 F 上所有 n 阶方阵构成的集合记为 $M_n(F)$.

线性方程组(2.1)可用矩阵表示如下:

$$\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n} & b_1\\ a_{21} & a_{22} & \cdots & a_{2n} & b_2\\ \vdots & \vdots & & \vdots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m\end{pmatrix},$$

它是一个 $m\times(n+1)$ 矩阵.第 i 行代表方程组(2.1)的第 i 个方程.对于 $1\leqslant j\leqslant n$,它的第 j 列表示的是方程组中 x_j 的系数.而最后一列由常数项组成.称它为线性方程组(2.1)的**增广矩阵**,而其中前 n 列组成的矩阵

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

称为方程组(2.1)的**系数矩阵**,它是由(2.1)的系数组成的.

方程组与矩阵可按上面的写法对应起来.而方程组的初等变换则对应着矩阵的初等行变换,其中的记号 ⓘ↔ⓙ 就表示交换矩阵的第 i, j 这两行,$c\cdot$ⓘ 表示用非 0 数 c 乘以矩阵的第 i 行,$k\cdot$ⓘ+ⓙ 表示把第 i 行的 k 倍加到第 j 行上.这样,例 2.1 中解方程组的过程可用矩阵表示如下:

$$\begin{pmatrix} 0 & 3 & 1 & 9 \\ 1 & -2 & 1 & 0 \\ 3 & 3 & -1 & 6 \end{pmatrix} \xrightarrow{①\leftrightarrow②} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 3 & 1 & 9 \\ 3 & 3 & -1 & 6 \end{pmatrix} \xrightarrow{-3\cdot ①+③} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 3 & 1 & 9 \\ 0 & 9 & -4 & 6 \end{pmatrix}$$

$$\xrightarrow{-3\cdot ②+③} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 3 & 1 & 9 \\ 0 & 0 & -7 & -21 \end{pmatrix} \xrightarrow{-\frac{1}{7}\cdot ③} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 3 & 1 & 9 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

$$\xrightarrow[-1\cdot ③+②]{-1\cdot ③+①} \begin{pmatrix} 1 & -2 & 0 & -3 \\ 0 & 3 & 0 & 6 \\ 0 & 0 & 1 & 3 \end{pmatrix} \xrightarrow{\frac{1}{3}\cdot ②} \begin{pmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

$$\xrightarrow{2\cdot ②+①} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

最后这个矩阵表示的线性方程组为 $\begin{cases} x_1 \quad\quad = 1, \\ \quad x_2 \quad = 2, \\ \quad\quad x_3 = 3. \end{cases}$ 由此可直接写出原方程组的解为(1,2,3).

例 2.2 解线性方程组

$$\begin{cases} x_1 - x_2 - 3x_3 + x_4 = 1, \\ x_1 - x_2 + 2x_3 - x_4 = 3, \\ 2x_1 - 2x_2 + 3x_3 - 4x_4 = 0, \\ 3x_1 - 3x_2 + 5x_3 \quad\quad = -1. \end{cases}$$

解

$$\begin{pmatrix}1&-1&-3&1&1\\1&-1&2&-1&3\\2&-2&3&-4&0\\3&-3&5&0&-1\end{pmatrix}\longrightarrow\begin{pmatrix}1&-1&-3&1&1\\0&0&5&-2&2\\0&0&9&-6&-2\\0&0&14&-3&-4\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&-1&-3&1&1\\0&0&5&-2&2\\0&0&45&-30&-10\\0&0&70&-15&-20\end{pmatrix}\longrightarrow\begin{pmatrix}1&-1&-3&1&1\\0&0&5&-2&2\\0&0&0&-12&-28\\0&0&0&13&-48\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&-1&-3&1&1\\0&0&5&-2&2\\0&0&0&-12&-28\\0&0&0&0&1\end{pmatrix}.$$

上面矩阵的最后一行对应的方程是

$$0\cdot x_1+0\cdot x_2+0\cdot x_3+0\cdot x_4=1.$$

不管 x_1,x_2,x_3,x_4 取何值,上式均不可能成立.所以原方程组无解.

例 2.3 解线性方程组

$$\begin{cases}x_1-x_2-3x_3+x_4=1,\\x_1-x_2+2x_3-x_4=3,\\4x_1-4x_2+3x_3-2x_4=10,\\2x_1-2x_2-11x_3+4x_4=0.\end{cases}$$

解

$$\begin{pmatrix}1&-1&-3&1&1\\1&-1&2&-1&3\\4&-4&3&-2&10\\2&-2&-11&4&0\end{pmatrix}\longrightarrow\begin{pmatrix}1&-1&-3&1&1\\0&0&5&-2&2\\0&0&15&-6&6\\0&0&-5&2&-2\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&-1&-3&1&1\\0&0&5&-2&2\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}\longrightarrow\begin{pmatrix}1&-1&-3&1&1\\0&0&1&-\frac{2}{5}&\frac{2}{5}\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & -1 & 0 & -\frac{1}{5} & \frac{11}{5}\\ 0 & 0 & 1 & -\frac{2}{5} & \frac{2}{5}\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}.$$

这时方程组化为

$$\begin{cases}x_1-x_2\quad-\frac{1}{5}x_4=\frac{11}{5},\\ \qquad\qquad x_3-\frac{2}{5}x_4=\frac{2}{5},\\ \qquad\qquad\qquad 0=0,\\ \qquad\qquad\qquad 0=0,\end{cases}$$

或写为

$$\begin{cases}x_1=x_2+\frac{1}{5}x_4+\frac{11}{5},\\ x_3=\quad\ \frac{2}{5}x_4+\frac{2}{5}.\end{cases}\tag{2.4}$$

可以看出,对于未知量 x_2,x_4 的任一组取值,都可以唯一定出 x_1,x_3 的值.称 x_1,x_3 为主未知量,x_2,x_4 为**自由未知量**.用自由未知量表示主未知量的(2.4)式称为方程组的一般解.任取自由未知量的一组值,由一般解可给出主未知量的值,从而得到方程组的一个解.由于自由未知量可以取数域 F 中任意数,它有无穷多组取法,所以原方程组有无穷多个解.

上面给出的求解线性方程组的方法也称为高斯(Gauss)消元法.

§ 2.2 线性方程组解的情况的判定

上一节作线性方程组的消元法中,我们按未知量 $x_1,x_2,\cdots,x_n$ 的次序依次在方程组的下面的方程中消去未知量.这样的方程组称为阶梯形方程组.而它的增广矩阵即称为阶梯形矩阵.

定义 2.2 一个矩阵若满足下面条件,称其为**阶梯形矩阵**.

(1) 矩阵若有零行(即元素全为 0 的行),则零行一定在矩阵的下方.

(2) 对于矩阵的每个非零行,从左看第 1 个非零元素称为此行的主元.矩阵下面行的主

元所在列一定在上面行的主元所在列的右端.

由定义可知,阶梯形矩阵中每个主元所在列中此主元下方的元素均为 0.

定义 2.3 一个阶梯形矩阵若满足下面条件,则称其为**简化的阶梯形矩阵**.

(1) 主元都是 1.

(2) 每个主元所在列中此主元上方的元素也都为 0.

命题 2.1 任一矩阵 $\boldsymbol{A}$ 都可以通过矩阵的初等行变换化为阶梯形矩阵.也可以通过初等行变换化成简化的阶梯形矩阵.

证明 对矩阵 $\boldsymbol{A}$ 的行数 m 作归纳.

$m=1$ 时,$\boldsymbol{A}$ 已经是阶梯形矩阵,结论显然成立.设结论对 $m-1$ 行矩阵成立,则对任一 m 行矩阵来说.若 $\boldsymbol{A}=\boldsymbol{O}$,即 $\boldsymbol{A}$ 的每个元素都为 0,则 $\boldsymbol{A}$ 是阶梯形矩阵.若 $\boldsymbol{A}\neq\boldsymbol{O}$,则 $\boldsymbol{A}$ 中一定有非零元素,取列指标最小的一个非零元素,设它在第 j 列(即 $\boldsymbol{A}$ 的第 1 到第 $j-1$ 列上的元素都为 0).由于可以任意交换两行,所以我们可以假设 $a_{1j}\neq 0$. 对任意第 i 行,$i=2,\cdots,m$,把 $\boldsymbol{A}$ 的第一行的 $-\dfrac{a_{ij}}{a_{1j}}$ 倍加到第 i 行上,设得到的矩阵为 $\boldsymbol{B}$,则矩阵 $\boldsymbol{B}$ 的第 j 列上除第 1 行元素外均为 0,即矩阵 $\boldsymbol{B}$ 形如

$$\boldsymbol{B}=\begin{pmatrix} 0 & \cdots & 0 & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ 0 & \cdots & 0 & 0 & b_{2,j+1} & \cdots & b_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & b_{m,j+1} & \cdots & b_{mn} \end{pmatrix}.$$

设 $\boldsymbol{B}$ 的后 $m-1$ 行作成的矩阵为 $\boldsymbol{B}_1$,则 $\boldsymbol{B}_1$ 为 $m-1$ 行矩阵.由归纳假设,$\boldsymbol{B}_1$ 通过初等行变换可以化为阶梯形矩阵.再加上矩阵 $\boldsymbol{B}$ 的第 1 行,得到的矩阵仍为阶梯形矩阵,即矩阵 $\boldsymbol{A}$ 通过初等行变换化成了阶梯形矩阵.由归纳法原理知,结论对任意 m 行矩阵都成立.

把 $\boldsymbol{A}$ 通过初等行变换化成阶梯形矩阵后,把每个非零行乘以适当的非零常数可把主元变成 1.再从最后一个主元开始,把它所在行的适当倍数加到其他非零行上可把它上方的元素变成 0. 以此类推,就把这个阶梯形矩阵化成简化的阶梯形矩阵了. □

注 上面命题的证明过程也给出了如何把一个矩阵通过初等行变换化成阶梯形矩阵以及简化的阶梯形矩阵的方法.

下面来讨论线性方程组的解的情况.线性方程组(2.1)有 n 个未知量,m 个方程.设它的增广矩阵为 $\overline{\boldsymbol{A}}$,$\overline{\boldsymbol{A}}$ 为 m 行 $n+1$ 列矩阵.对矩阵 $\overline{\boldsymbol{A}}$ 作初等行变换化成阶梯形矩阵 $\boldsymbol{B}$,则矩阵 $\boldsymbol{B}$ 对应的线性方程组与方程组(2.1)同解.

(1) 若矩阵 $\boldsymbol{B}$ 有一个主元在最后一列,即矩阵 $\boldsymbol{B}$ 中除全零行外,倒数第 1 个非零行形如

$$(0,0,\cdots,0,0,d),$$

其中 $d\neq 0$.这一行对应的方程为

$$0\cdot x_1+0\cdot x_2+\cdots+0\cdot x_{n-1}+0\cdot x_n=d,$$

显然,不论 $x_1,x_2,\cdots,x_n$ 取何值,上式均不成立.所以这个方程无解.也即原方程组(2.1)无解.

(2) 若矩阵 $\boldsymbol{B}$ 的主元都不在最后一列.设这时矩阵 $\boldsymbol{B}$ 有 r 个非零行,也即有 r 个主元,这时一定有 $r\leqslant n$.(因为每一列上最多有一个主元,而主元又在前 n 列上).

① 若 $r=n$,这时前 n 列上每一列都有一个主元.即这时矩阵 $\boldsymbol{B}$ 形如

$$\boldsymbol{B}=\begin{pmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1n} & d_1 \\ 0 & b_{22} & b_{23} & \cdots & b_{2n} & d_2 \\ 0 & 0 & b_{33} & \cdots & b_{3n} & d_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & b_{nn} & d_n \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

其中 $b_{ii}\neq 0,1\leqslant i\leqslant n$.继续对矩阵 $\boldsymbol{B}$ 作初等行变换化为简化的阶梯形矩阵

$$\boldsymbol{C}=\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & 0 & \cdots & 0 & c_2 \\ 0 & 0 & 1 & \cdots & 0 & c_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & c_n \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

可得方程组有唯一解$(c_1,c_2,\cdots,c_n)$.

② 若 $r<n$,由于调换矩阵的列仅仅是交换未知量的次序,故不妨设这时主元所在的列为前 r 列.即 $\boldsymbol{B}$ 形如

$$\boldsymbol{B}=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1r} & b_{1,r+1} & \cdots & b_{1n} & d_1 \\ 0 & b_{22} & \cdots & b_{2r} & b_{2,r+1} & \cdots & b_{2n} & d_2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & b_{rr} & b_{r,r+1} & \cdots & b_{rn} & d_r \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

其中 $b_{ii}\neq 0,1\leqslant i\leqslant r$.依然继续对矩阵 $\boldsymbol{B}$ 作初等行变换化为简化的阶梯形矩阵

$$
\boldsymbol{C}=\begin{pmatrix}
1 & 0 & \cdots & 0 & c_{1,r+1} & \cdots & c_{1n} & c_1\\
0 & 1 & \cdots & 0 & c_{2,r+1} & \cdots & c_{2n} & c_2\\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & \cdots & 1 & c_{r,r+1} & \cdots & c_{rn} & c_r\\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0\\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0
\end{pmatrix},
$$

这时方程组可写为

$$
\begin{cases}
x_1=-c_{1,r+1}x_{r+1}-\cdots-c_{1n}x_n+c_1,\\
x_2=-c_{2,r+1}x_{r+1}-\cdots-c_{2n}x_n+c_2,\\
\cdots\cdots\cdots\cdots\\
x_r=-c_{r,r+1}x_{r+1}-\cdots-c_{rn}x_n+c_r.
\end{cases}
\tag{2.5}
$$

对于未知量 $x_{r+1},\cdots,x_n$ 的任一组取值,由(2.5)可以得到未知量 $x_1,\cdots,x_r$ 的值.所以这时原方程组有无穷多个解.这时我们也称未知量 $x_{r+1},\cdots,x_n$ 为自由未知量,$x_1,\cdots,x_r$ 为主未知量.用自由未知量表示主未知量的关系式(2.5)称为方程组的一般解.

综合上面的讨论.可得到下面的结论.

定理 2.1 设含 n 个未知量的线性方程组的增广矩阵为 $\overline{\boldsymbol{A}}$,对 $\overline{\boldsymbol{A}}$ 作初等行变换化为阶梯形矩阵 $\boldsymbol{B}$.若 $\boldsymbol{B}$ 有一个主元在最后一列,则方程组无解.若 $\boldsymbol{B}$ 的主元都不在最后一列.则方程组有解.进一步地,若这时 $\boldsymbol{B}$ 的非零行个数 $r=n$,则方程组有唯一解.若 $r<n$,则方程组有无穷多个解.

注 当方程组有解时.我们要继续把阶梯形矩阵 $\boldsymbol{B}$ 用初等行变换化成简化的阶梯形矩阵.由此可以直接写出原方程组的解(唯一解或一般解).

§2.3 齐次线性方程组

常数项全为零的线性方程组称为齐次线性方程组.这时(0,0,⋯,0)一定是它的解,称为零解.或者因为齐次线性方程组的增广矩阵的最后一列的元素全为0,所以作初等行变换后最后一列的元素也全为0,主元不可能在最后一列,由上面定理2.1知,齐次线性方程组一定有解.

由定理2.1知,若齐次线性方程组有非零解,则它就有无穷多个解,反之亦然.

对于齐次线性方程组.我们最关心的是它有没有非零解.下面的定理给出一个充分条件.

定理 2.2 若齐次线性方程组的方程个数 m 小于未知量个数 n，则它一定有非零解.

证明 对此方程组的增广矩阵作初等行变换化为阶梯形矩阵，这个阶梯形矩阵的非零行个数 r 一定小于或等于 m，故 $r\leqslant m<n$，即 $r<n$. 由定理 2.1 知方程组有无穷多个解. □

另外，在具体求解齐次线性方程组时，由于它的增广矩阵的最后一列元素都是 0，作初等行变换也不改变这个事实，所以只需要应用它的系数矩阵就够了.

例 2.4 求解齐次线性方程组

$$\begin{cases} 2x_1 - x_2 + 5x_3 - 3x_4 = 0, \\ x_1 - 5x_2 + 3x_3 + 2x_4 = 0, \\ 3x_1 - 4x_2 + 7x_3 - x_4 = 0, \\ 9x_1 - 7x_2 + 15x_3 + 4x_4 = 0. \end{cases}$$

解
$$A = \begin{pmatrix} 2 & -1 & 5 & -3 \\ 1 & -5 & 3 & 2 \\ 3 & -4 & 7 & -1 \\ 9 & -7 & 15 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -5 & 3 & 2 \\ 2 & -1 & 5 & -3 \\ 3 & -4 & 7 & -1 \\ 9 & -7 & 15 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -5 & 3 & 2 \\ 0 & 9 & -1 & -7 \\ 0 & 11 & -2 & -7 \\ 0 & 38 & -12 & -14 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & -5 & 3 & 2 \\ 0 & 9 & -1 & -7 \\ 0 & 11 & -2 & -7 \\ 0 & 19 & -6 & -7 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -5 & 3 & 2 \\ 0 & 9 & -1 & -7 \\ 0 & 11 & -2 & -7 \\ 0 & 1 & -4 & 7 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -5 & 3 & 2 \\ 0 & 1 & -4 & 7 \\ 0 & 11 & -2 & -7 \\ 0 & 9 & -1 & -7 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & -5 & 3 & 2 \\ 0 & 1 & -4 & 7 \\ 0 & 0 & 42 & -84 \\ 0 & 0 & 35 & -70 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -5 & 3 & 2 \\ 0 & 1 & -4 & 7 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -5 & 0 & 8 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

所以原方程组有无穷多个解. 它的一般解为

$$\begin{cases} x_1 = -3x_4, \\ x_2 = x_4, \\ x_3 = 2x_4, \end{cases}$$

其中 x_4 为自由未知量.

习 题 二

1. 解下列线性方程组:

(1) $\begin{cases} x_1-3x_2-2x_3=3, \\ -2x_1+x_2-4x_3=-9, \\ -x_1+4x_2-x_3=-7; \end{cases}$

(2) $\begin{cases} x_1+3x_2+2x_3=1, \\ 2x_1+5x_2+5x_3=7, \\ 3x_1+7x_2+x_3=-8, \\ -x_1-4x_2+x_3=10; \end{cases}$

(3) $\begin{cases} x_1-3x_2-2x_3-x_4=6, \\ 3x_1-8x_2+x_3+5x_4=0, \\ -2x_1+x_2-4x_3+x_4=-12, \\ -x_1+4x_2-x_3-3x_4=2; \end{cases}$

(4) $\begin{cases} x_1+3x_2-7x_3=-8, \\ 2x_1+5x_2+4x_3=4, \\ -3x_1-7x_2-2x_3=-3, \\ x_1+4x_2-12x_3=-15; \end{cases}$

(5) $\begin{cases} x_1-2x_2+3x_3-4x_4=4, \\ x_1+x_2-x_3+x_4=-11, \\ x_1+3x_2+x_4=1, \\ -7x_2+3x_3+x_4=-3; \end{cases}$

(6) $\begin{cases} 2x_1-3x_2+x_3+5x_4=6, \\ -3x_1+x_2+2x_3-4x_4=5, \\ -x_1-2x_2+3x_3+x_4=-2. \end{cases}$

2. 当 a 为何值时,下述线性方程组有解?当有解时,求出它的所有解.

$$\begin{cases} x_1-4x_2+2x_3=-1, \\ -x_1+11x_2-x_3=3, \\ 3x_1-5x_2+7x_3=a. \end{cases}$$

3. 当 a 为何值时，下述线性方程组有唯一解？无解？

$$\begin{cases} x_1+x_2+x_3=3, \\ x_1+2x_2-ax_3=9, \\ 2x_1-x_2+3x_3=6. \end{cases}$$

4. (1) 下述线性方程组有多少个解？

$$\begin{cases} x+y=1, \\ x-3y=-1, \\ 10x-4y=3; \end{cases}$$

(2) 在平面直角坐标系中画出第(1)小题中各个方程表示的图形.

5. 当 a 为何值时，下述线性方程组有解？当有解时，求它的所有解.

$$\begin{cases} x_1+x_2+x_3+x_4=-7, \\ x_1+3x_3-x_4=8, \\ x_1+2x_2-x_3+x_4=2a+2, \\ 3x_1+3x_2+3x_3+2x_4=-11, \\ 2x_1+2x_2+2x_3+x_4=2a. \end{cases}$$

6. 下列齐次线性方程组有无非零解？若有非零解，求出它的一般解.

(1) $\begin{cases} 3x_1-5x_2+x_3-2x_4=0, \\ 2x_1+3x_2-5x_3+x_4=0, \\ -x_1+7x_2-4x_3+3x_4=0, \\ 4x_1+15x_2-7x_3+9x_4=0; \end{cases}$　(2) $\begin{cases} 5x_1-2x_2+4x_3-3x_4=0, \\ -3x_1+5x_2-x_3+2x_4=0, \\ x_1-3x_2+2x_3+x_4=0. \end{cases}$

第三章　行列式

§3.1　n 元排列

第零章中我们已见到 2 阶及 3 阶行列式,本章将定义任意的 n 阶行列式.为此,先来看 n 元排列.

定义 3.1　n 个数 $1,2,\cdots,n$ 的一个排列是由这 n 个数组成的一个有序数串,也即这 n 个元素所组成的一个**全排列**.这样的一个排列也称为一个 **n 元排列**.

显然,n 元排列的个数为 $n\cdot(n-1)\cdots2\cdot1=n!$.

定义 3.2　在一个 n 元排列 $i_1i_2\cdots i_n$ 中.若较大的数排在较小的数的前面,即当 $t>s$ 时有 $i_s>i_t$,则称这一对数 i_si_t 构成一个**逆序**.此排列 $i_1i_2\cdots i_n$ 中的所有逆序的个数,称为它的**逆序数**,记为 $\tau(i_1i_2\cdots i_n)$.

例如,在 7 元排列 4765132 中,41,43,42,76,75,71,73,72,65,61,63,62,51,53,52,32 为所有的逆序.所以 $\tau(4765132)=16$.

求一个排列的逆序数,可以从第一个数字开始,依次检查它是否与其后面的数组成逆序.接着看第二个数字.以此类推就可以求出所有逆序,也就求出了它的逆序数了.

例 3.1　$\tau(12\cdots n)=0$,而

$$\tau(n(n-1)\cdots21)=(n-1)+(n-2)+\cdots+1=\frac{n}{2}(n-1).$$

定义 3.3　逆序数为偶数的排列称为**偶排列**.逆序数为奇数的排列称为**奇排列**.

故 7 元排列 4765132 为偶排列,n 元排列 $12\cdots n$ 为偶排列.而对 n 元排列 $n(n-1)\cdots21$ 来说,若 $n\equiv0,1\pmod 4$,则它为偶排列,若 $n\equiv2,3\pmod 4$,则它为奇排列.

下面用 P_n 表示 $1,2,\cdots,n$ 所构成的所有 n 元排列组成的集合.下面定义 P_n 到其自身的一个映射.

定义 3.4　对于任意取定的两个互异元素 $i,j\in\{1,2,\cdots,n\}$.对换 (i,j) 表示把 P_n 中每个排列的 i,j 这两个元素互换位置而其余元素不变所得到的 P_n 到 P_n 的映射,即

$$(i,j): P_n \longrightarrow P_n$$

$$\cdots i \cdots j \cdots \longmapsto \cdots j \cdots i \cdots$$

例如对换(2,5)把 7 元排列 4765132 变成 4762135,把 1235746 变成 1532746.容易得到 $\tau(4765132)=16$ 而 $\tau(4762135)=3+5+4+1=13$,$\tau(1235746)=1+2=3$,而 $\tau(1532746)=3+1+2=6$,即对换把偶排列变成了奇排列,把奇排列变成了偶排列.一般地,这个结论也成立.

定理 3.1 对换改变排列的奇偶性.

证明 考察对换(i,j).

(i) 设 i,j 为排列中相邻的两个数,即排列形如$\cdots ij\cdots$,称它为排列(Ⅰ).做对换(i,j)后,排列(Ⅰ)变成排列(Ⅱ)$\cdots ji\cdots$.

在排列(Ⅱ)中,i,j 以外的数之间是否构成逆序与排列(Ⅰ)相同,i,j 以外的数与 i,j 是否构成逆序与排列(Ⅰ)也相同,唯一不同的是 i,j 这两个数是否构成逆序.所以当 $i<j$ 时,在排列(Ⅰ)中 ij 不是逆序,而在排列(Ⅱ)中 ji 为逆序,这时 $\tau(\text{Ⅱ})=\tau(\text{Ⅰ})+1$.而当 $i>j$ 时,在排列(Ⅰ)中 ij 是逆序,到了排列(Ⅱ)中 ji 就不是逆序了.这时 $\tau(\text{Ⅱ})=\tau(\text{Ⅰ})-1$,即排列(Ⅱ)与排列(Ⅰ)的逆序数相差 1,它们的奇偶性不同.

(ii) 若排列中 i,j 这两个数不相邻.设它们之间有 $s(\geqslant 1)$ 个数 $i_1,i_2,\cdots,i_s$.设排列(Ⅰ)形如$\cdots ii_1i_2\cdots i_sj\cdots$,作对换$(i,j)$后,变为排列(Ⅱ)$\cdots ji_1i_2\cdots i_si\cdots$.这个对换$(i,j)$可用下面 $2s+1$ 个相邻两个数的对换来完成

$$\cdots ii_1i_2\cdots i_sj\cdots \xrightarrow{(i,i_1)} \cdots i_1ii_2\cdots i_sj\cdots \xrightarrow{(i,i_2)} \cdots i_1i_2i\cdots i_sj\cdots \to \cdots \xrightarrow{(i,i_s)} \cdots i_1i_2\cdots i_sij\cdots$$

$$\xrightarrow{(i,j)} \cdots i_1i_2\cdots i_sji\cdots \xrightarrow{(i_s,j)} \cdots i_1i_2\cdots ji_si\cdots \to \cdots \xrightarrow{(i_2,j)} \cdots i_1ji_2\cdots i_si\cdots \xrightarrow{(i_1,j)} \cdots ji_1i_2\cdots i_si\cdots.$$

由(i)知每做一次相邻两个数的对换都改变排列的奇偶性,所以共进行了奇数次的相邻两个数的对换.最后也就改变了排列的奇偶性. □

性质 3.1 n 元排列中奇、偶排列都有$\frac{1}{2}n!$个.

证明 用 M 表示 n 元奇排列的集合,N 表示 n 元偶排列的集合,则对换(1,2)是从 M 到 N 的一个双射,所以 $|M|=|N|$.又 $|M|+|N|=n!$,所以 $|M|=|N|=\frac{1}{2}n!$. □

性质 3.2 任一 n 元排列可通过一系列对换变成 $12\cdots n$,且所作的对换的个数与该排列有相同的奇偶性.

证明 先证任一 n 元排列都可通过一系列对换变成 $12\cdots n$.对 n 作归纳.$n=1$ 时结论显然成立.设结论对任意 $n-1$ 元排列成立,则对任一 n 元排列 $i_1i_2\cdots i_n$,若 $i_n=n$,则 n 已在此排列的最后一个位置上.若 $i_n\neq n$,则作对换(i_n,n)可把数字 n 换到排列的最后一个位置上,即为 $j_1j_2\cdots j_{n-1}n$,其中 $j_1j_2\cdots j_{n-1}$ 为一个 $n-1$ 元排列.由归纳假设,$j_1j_2\cdots j_{n-1}$ 可通过一系列对换变成 $12\cdots(n-1)$.从而 $i_1i_2\cdots i_n$ 就通过这些对换和对换(i_n,n)变成了 $12\cdots n$.

再证明所做对换的个数 s 与 $\tau(i_1i_2\cdots i_n)$ 有相同的奇偶性.由于每做一次对换都改变了排

列的奇偶性.而排列 $12\cdots n$ 为偶排列,即 $\tau(i_1i_2\cdots i_n)$ 的奇偶性变换 s 次后变成了偶数.所以 $\tau(i_1i_2\cdots i_n)$ 与 s 有相同的奇偶性. □

例如:$4765132\xrightarrow{(7,2)}4265137\xrightarrow{(6,3)}4235167\xrightarrow{(5,1)}4231567\xrightarrow{(4,1)}1234567$,所做对换的次数 4 与 $\tau(4765132)=16$ 都是偶数.

§3.2 n阶行列式

定义 3.5 设 n 阶矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

则 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|$,或记为 $\det(\boldsymbol{A})$,定义为

$$|\boldsymbol{A}|=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=\sum_{j_1j_2\cdots j_n\in P_n}(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n},$$

任一 n 阶矩阵的行列式也称为一个 n 阶行列式.

由定义可以看出,n 阶行列式是 $n!$ 项的代数和.每一项都是取自不同行、不同列上的 n 个元素的乘积,而这一项所带的符号是由当把乘积的这 n 个元素的行指标按自然次序 $1,2,\cdots,n$ 排好后,它的列指标排列 $j_1j_2\cdots j_n$ 的奇偶性所确定的.偶排列对应的是正号,奇排列对应的是负号.

例 3.2 1 阶行列式 $|a_{11}|=a_{11}$,2 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=\sum_{j_1j_2\in P_2}(-1)^{\tau(j_1j_2)}a_{1j_1}a_{2j_2}=a_{11}a_{22}-a_{12}a_{21},$$

这等于主对角线上两个元素的乘积减去副对角线上两个元素的乘积.

3 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=\sum_{j_1j_2j_3\in P_3}(-1)^{\tau(j_1j_2j_3)}a_{1j_1}a_{2j_2}a_{3j_3},$$

3 元排列共 3！=6 个.其中偶排列为 123,231,312,奇排列为 132,213,321,所以

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31},$$

即为平行于主对角线的 3 条线上元素的乘积之和再减去平行于副对角线的 3 条线上元素乘积的和.

例 3.3 计算 4 阶行列式 $\begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 5 & 0 & 0 \\ 6 & 0 & 0 & 0 \end{vmatrix}$.

解 原式 $= \sum\limits_{j_1j_2j_3j_4} (-1)^{\tau(j_1j_2j_3j_4)} a_{1j_1}a_{2j_2}a_{3j_3}a_{4j_4}$.

由于当 $j_1 \neq 4$ 时 $a_{1j_1}=0$,这时包含 a_{1j_1} 的项就为 0.所以要求值,只需考察 $j_1=4$ 的情形.同理也只需考察 $j_2=3, j_3=2, j_4=1$ 的情形,故

$$\text{原式} = (-1)^{\tau(4321)} 1 \cdot 2 \cdot 5 \cdot 6 = 60.$$

例 3.4 主对角线下方的元素都为 0 的方阵称为**上三角矩阵**.它的行列式称为**上三角行列式**,下面计算上三角行列式.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = \sum_{j_1j_2\cdots j_n \in P_n} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}a_{2j_2}\cdots a_{nj_n}.$$

当 $j_n \neq n$ 时 $a_{nj_n}=0$,所以含有这样的 a_{nj_n} 的项为 0,故为求行列式只需设 $j_n=n$.这时 $j_1j_2\cdots j_{n-1}$ 是 $1,2,\cdots,n-1$ 上的一个 $n-1$ 元排列,即 $j_{n-1} \neq n$.而当 $j_{n-1}<n-1$ 时,有 $a_{n-1,j_{n-1}}=0$.所以只需设 $j_{n-1}=n-1$.以此类推,上述行列式除一可能项 $a_{11}a_{22}\cdots a_{nn}$ 外,其余 $n!-1$ 项中一定含有 0 这个元素,即这样的项为 0,所以

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = (-1)^{\tau(12\cdots n)} a_{11}a_{22}\cdots a_{nn} = a_{11}a_{22}\cdots a_{nn}.$$

即上三角行列式等于它的主对角线上元素的乘积.

行列式展开式中的每一项都是取自不同行不同列的 n 个元素的乘积,如果它的行指标不按自然顺序 $1,2,\cdots,n$ 排列时,此项所带的符号如何确定？因为数的乘积具有交换性,所以我们可以把此项中的元素任意交换次序,而每交换一次次序,则相当于此项 n 个元素的行

指标作了一次对换，列指标也作了一次对换，所以它们的行指标排列与列指标排列的逆序数之和的奇偶性不变.设这样的一项记为 $a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}$，则我们交换这 n 个元素 $a_{i_1j_1},a_{i_2j_2},\cdots,a_{i_nj_n}$ 在乘积中的前后次序变成 $a_{i_1'j_1'}a_{i_2'j_2'}\cdots a_{i_n'j_n'}$时 $\tau(i_1i_2\cdots i_n)+\tau(j_1j_2\cdots j_n)$ 与 $\tau(i_1'i_2'\cdots i_n')+\tau(j_1'j_2'\cdots j_n')$ 有相同的奇偶性.所以当把这项变成 $a_{1k_1}a_{2k_2}\cdots a_{nk_n}$时，有 $\tau(i_1i_2\cdots i_n)+\tau(j_1j_2\cdots j_n)$ 与 $\tau(12\cdots n)+\tau(k_1k_2\cdots k_n)=\tau(k_1k_2\cdots k_n)$ 有相同的奇偶性.由于 $a_{1k_1}a_{2k_2}\cdots a_{nk_n}$ 在行列式展开式中所带符号为 $(-1)^{\tau(k_1k_2\cdots k_n)}$，所以 $a_{i_1j_1}a_{i_2j_2}\cdots a_{i_nj_n}$ 在行列式展开式中所带符号为 $(-1)^{\tau(i_1i_2\cdots i_n)+\tau(j_1j_2\cdots j_n)}$.特别地，当我们把展开式中的各项都按列指标为自然顺序排列时，即为 $a_{i_11}a_{i_22}\cdots a_{i_nn}$，则此项所带的符号为 $(-1)^{\tau(i_1i_2\cdots i_n)}$，所以我们可给出行列式定义的另一个描述为

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{i_1i_2\cdots i_n\in P_n}(-1)^{\tau(i_1i_2\cdots i_n)}a_{i_11}a_{i_22}\cdots a_{i_nn}.$$

§3.3 行列式的性质

性质 3.3 行列互换，行列式的值不变，即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix}.$$

对任一 $m\times n$ 矩阵 $\boldsymbol{A}$，将它的行列互换，即把 $\boldsymbol{A}$ 的第 i 行变成第 i 列，$1\leqslant i\leqslant m$，所得到的 $n\times m$ 矩阵，称为矩阵 $\boldsymbol{A}$ 的**转置**，记为 $\boldsymbol{A}^{\mathrm{T}}$ 或 $\boldsymbol{A}'$. 所以矩阵 $\boldsymbol{A}$ 的第 i 行第 j 列元素是 $\boldsymbol{A}^{\mathrm{T}}$ 的第 j 行第 i 列元素.上述行列式的性质，即是证 n 阶方阵 $\boldsymbol{A}$ 与它的转置 $\boldsymbol{A}^{\mathrm{T}}$ 有相同的行列式.

证明 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$，$\boldsymbol{A}^{\mathrm{T}}=(b_{ij})_{n\times n}$，则 $b_{ij}=a_{ji}$，所以

$$\begin{aligned} |\boldsymbol{A}^{\mathrm{T}}| = |(b_{ij})_{n\times n}| &= \sum_{i_1i_2\cdots i_n\in P_n}(-1)^{\tau(i_1i_2\cdots i_n)}b_{i_11}b_{i_22}\cdots b_{i_nn} \\ &= \sum_{i_1i_2\cdots i_n\in P_n}(-1)^{\tau(i_1i_2\cdots i_n)}a_{1i_1}a_{2i_2}\cdots a_{ni_n} = |\boldsymbol{A}|. \end{aligned}$$ □

主对角线上方全为 0 的方阵称为**下三角矩阵**，其行列式也称为**下三角行列式**.由性质 3.3 知

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ 0 & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

所以下三角行列式也等于它主对角线上元素之积.

性质 3.3 表明行列式中对行成立的性质,对列来说也照样成立,所以下面行列式的性质都仅对行的情形加以说明.

性质 3.4 行列式中某一行中元素的公因子可以提到行列式外,即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

证明

$$\begin{aligned} \text{左端} &= \sum_{j_1j_2\cdots j_n \in P_n} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}\cdots(ka_{ij_i})\cdots a_{nj_n} \\ &= k\sum_{j_1j_2\cdots j_n \in P_n} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}\cdots a_{ij_i}\cdots a_{nj_n} = \text{右端}. \end{aligned}$$ □

性质 3.4 也表明用一个数 k 去乘一个行列式等于用 k 去乘此行列式的某一行.

性质 3.5

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1}+c_{i1} & b_{i2}+c_{i2} & \cdots & b_{in}+c_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ c_{i1} & c_{i2} & \cdots & c_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

证明

$$\begin{aligned} \text{左端} &= \sum_{j_1j_2\cdots j_n \in P_n} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}\cdots(b_{ij_i}+c_{ij_i})\cdots a_{nj_n} \\ &= \sum_{j_1j_2\cdots j_n \in P_n} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}\cdots b_{ij_i}\cdots a_{nj_n} + \sum_{j_1j_2\cdots j_n \in P_n} (-1)^{\tau(j_1j_2\cdots j_n)} a_{1j_1}\cdots c_{ij_i}\cdots a_{nj_n} \\ &= \text{右端}. \end{aligned}$$ □

性质 3.5 表明,行列式中若有某一行是两组数之和,则此行列式等于两个行列式的和,这两个行列式的这一行分别是这两组数的第一组数和第二组数,而其余各行均保持不变.

例 3.5 $\begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{vmatrix}$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}.$$

性质 3.6 交换行列式的两行,行列式的值变号,即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

证明 记右端行列式第(i,j)位置的元素为b_{ij}.即对任意$1\leqslant t\leqslant n$,当$k\neq i,j$时$b_{kt}=a_{kt}$,而$b_{it}=a_{jt}$,$b_{jt}=a_{it}$,则

$$\begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & & \vdots \\ b_{j1} & b_{j2} & \cdots & b_{jn} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix} = \sum_{k_1\cdots k_i\cdots k_j\cdots k_n\in P_n} (-1)^{\tau(k_1\cdots k_i\cdots k_j\cdots k_n)} b_{1k_1}\cdots b_{ik_i}\cdots b_{jk_j}\cdots b_{nk_n}$$

$$= \sum_{k_1\cdots k_i\cdots k_j\cdots k_n\in P_n} (-1)^{\tau(k_1\cdots k_i\cdots k_j\cdots k_n)} a_{1k_1}\cdots a_{jk_i}\cdots a_{ik_j}\cdots a_{nk_n}$$

$$= - \sum_{k_1\cdots k_j\cdots k_i\cdots k_n\in P_n} (-1)^{\tau(k_1\cdots k_j\cdots k_i\cdots k_n)} a_{1k_1}\cdots a_{ik_j}\cdots a_{jk_i}\cdots a_{nk_n}$$

$$= - \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$ □

推论 3.1 两行相同,则行列式值为 0.

证明 交换这相同的两行,则得到的是同一个行列式,而由性质 3.6 知,这时值又变号,即此行列式与它的相反数相等,故其值为 0. □

推论 3.2 两行成比例,则行列式值为 0.

证明 由性质 3.4 和推论 3.1 立得. □

性质 3.7 把行列式某一行的倍数加到另一行上,值不变,即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1}+ka_{i1} & a_{j2}+ka_{j2} & \cdots & a_{jn}+ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

证明 由性质 3.5 和推论 3.2 有

$$\text{左端} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \text{右端}.$$ □

上面给出了当对矩阵施行初等行变换时,其行列式的变化规律.每一个矩阵都可以通过矩阵的初等行变换化成阶梯形,而阶梯形矩阵又是上三角形的,由此我们便可以求出行列式的值.

例 3.6 计算行列式

$$\begin{vmatrix} 2 & 1 & -5 & 8 \\ 1 & -3 & 0 & 9 \\ 0 & 2 & -1 & -5 \\ 1 & 4 & -7 & 0 \end{vmatrix}.$$

解

$$\text{原式}=-\begin{vmatrix}1&-3&0&9\\2&1&-5&8\\0&2&-1&-5\\1&4&-7&0\end{vmatrix}=-\begin{vmatrix}1&-3&0&9\\0&7&-5&-10\\0&2&-1&-5\\0&7&-7&-9\end{vmatrix}=-\begin{vmatrix}1&-3&0&9\\0&1&-2&5\\0&2&-1&-5\\0&7&-7&-9\end{vmatrix}$$

$$=-\begin{vmatrix}1&-3&0&9\\0&1&-2&5\\0&0&3&-15\\0&0&7&-44\end{vmatrix}=-\begin{vmatrix}1&-3&0&9\\0&1&-2&5\\0&0&3&-15\\0&0&1&-14\end{vmatrix}=\begin{vmatrix}1&-3&0&9\\0&1&-2&5\\0&0&1&-14\\0&0&3&-15\end{vmatrix}$$

$$=\begin{vmatrix}1&-3&0&9\\0&1&-2&5\\0&0&1&-14\\0&0&0&27\end{vmatrix}=1\times1\times1\times27=27.$$

例 3.7 计算下列 n 阶行列式

(1) $\begin{vmatrix}a&b&b&\cdots&b\\b&a&b&\cdots&b\\b&b&a&\cdots&b\\\vdots&\vdots&\vdots&&\vdots\\b&b&b&\cdots&a\end{vmatrix}$； (2) $\begin{vmatrix}a_1&a_2&a_3&\cdots&a_n\\b_2&c_2&0&\cdots&0\\b_3&0&c_3&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\b_n&0&0&\cdots&c_n\end{vmatrix}$，$c_i\neq0,2\leqslant i\leqslant n$.

解 (1) 原式 $=\begin{vmatrix}a+(n-1)b&b&b&\cdots&b\\a+(n-1)b&a&b&\cdots&b\\a+(n-1)b&b&a&\cdots&b\\\vdots&\vdots&\vdots&&\vdots\\a+(n-1)b&b&b&\cdots&a\end{vmatrix}$

$$=[a+(n-1)b]\cdot\begin{vmatrix}1&b&b&\cdots&b\\1&a&b&\cdots&b\\1&b&a&\cdots&b\\\vdots&\vdots&\vdots&&\vdots\\1&b&b&\cdots&a\end{vmatrix}$$

$$=[a+(n-1)b]\cdot\begin{vmatrix}1 & b & b & \cdots & b\\ 0 & a-b & 0 & \cdots & 0\\ 0 & 0 & a-b & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & a-b\end{vmatrix}$$

$$=[a+(n-1)b]\cdot(a-b)^{n-1}.$$

(2) 原式 $=\begin{vmatrix}a_1-\frac{a_2}{c_2}b_2-\frac{a_3}{c_3}b_3-\cdots-\frac{a_n}{c_n}b_n & 0 & 0 & \cdots & 0\\ b_2 & c_2 & 0 & \cdots & 0\\ b_3 & 0 & c_3 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ b_n & 0 & 0 & \cdots & c_n\end{vmatrix}$

$$=\left(a_1-\frac{a_2}{c_2}b_2-\frac{a_3}{c_3}b_3-\cdots-\frac{a_n}{c_n}b_n\right)c_2c_3\cdots c_n.$$

§3.4 行列式的展开公式

定义 3.6 在 n 阶行列式 $|a_{ij}|_{n\times n}$ 中,去掉其中的第 i 行和第 j 列元素,剩下的元素按原来的位置所构成的 $n-1$ 阶行列式

$$M_{ij}=\begin{vmatrix}a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n}\\ \vdots & & \vdots & \vdots & & \vdots\\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n}\\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n}\\ \vdots & & \vdots & \vdots & & \vdots\\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn}\end{vmatrix}$$

称为元素 a_{ij} 的**余子式**,而 $A_{ij}=(-1)^{i+j}M_{ij}$ 称为元素 a_{ij} 的**代数余子式**.

例如,行列式

$$\begin{vmatrix}1 & 2 & 1 & 4\\ 0 & -1 & 2 & 1\\ 5 & 3 & 6 & 0\\ 1 & 0 & 3 & 3\end{vmatrix}$$

中第 3 行第 2 列元素 3 的余子式为

$$\begin{vmatrix} 1 & 1 & 4 \\ 0 & 2 & 1 \\ 1 & 3 & 3 \end{vmatrix} = -4,$$

代数余子式为

$$(-1)^{3+2}\begin{vmatrix} 1 & 1 & 4 \\ 0 & 2 & 1 \\ 1 & 3 & 3 \end{vmatrix} = -(-4) = 4.$$

由定义知,元素 a_{ij}的余子式以及代数余子式与第 i 行和第 j 列上的元素无关.

命题 3.1
$$\begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} = a_{ij}A_{ij},$$

其中 A_{ij}为元素 a_{ij}的代数余子式.

证明 先证 $i=j=n$ 的情形.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & & & \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{vmatrix} = \sum_{j_1j_2\cdots j_{n-1}j_n \in P_n} (-1)^{\tau(j_1j_2\cdots j_{n-1}j_n)} a_{1j_1}a_{2j_2}\cdots a_{n-1,j_{n-1}}a_{nj_n}$$

$$= a_{nn}\sum_{j_1j_2\cdots j_{n-1} \in P_{n-1}} (-1)^{\tau(j_1j_2\cdots j_{n-1})} a_{1j_1}a_{2j_2}\cdots a_{n-1,j_{n-1}}$$

$$= a_{nn}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} \\ \vdots & \vdots & & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} \end{vmatrix} = a_{nn}M_{nn} = a_{nn}A_{nn}.$$

再看一般的情形,把行列式中的第 i 行与第 $i+1$ 行互换,再与第 $i+2$ 行互换,…,与第 n 行互换,一共做了 $n-i$ 次两行的对换,之后把第 j 列与第 $j+1$ 列对换,再与第 $j+2$ 列对换,一直下去,最后与第 n 列对换,一共做了 $n-j$ 次两列的对换,就把第 i 行换到第 n 行,第 j 列换到第 n

列,而其余行、列的前后次序不变,所以

$$\begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

$$= (-1)^{(n-i)+(n-j)} \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} & a_{1j} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} & a_{i-1,j} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} & a_{i+1,j} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} & a_{nj} \\ 0 & \cdots & 0 & 0 & \cdots & 0 & a_{ij} \end{vmatrix}$$

$$= (-1)^{2n-(i+j)} a_{ij} \cdot M_{ij} = a_{ij}A_{ij}.$$ □

定理 3.2 设 $A=(a_{ij})_{n\times n}$,则 n 阶行列式 $|A|$ 等于它的任意行的各元素与其相应的代数余子式的乘积之和,即对任意 $1\leqslant i\leqslant n$,有

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{j=1}^{n} a_{ij}A_{ij}.$$

证明 由行列式性质 3.5,将 $|A|$ 的第 i 行元素中的每一个都写成 n 个数之和

$$a_{i1}+0+\cdots+0,\quad 0+a_{i2}+\cdots+0,\quad \cdots,\quad 0+0+\cdots+a_{in},$$

则 $|A|$ 可以写成 n 个行列式之和,即

$$|A| = \sum_{j=1}^{n} \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & & & & & \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & a_{ij} & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

再由命题 3.1 即得

$$|\boldsymbol{A}| = \sum_{j=1}^{n} a_{ij}A_{ij}. \qquad \square$$

由于行列式的行、列具有同样的性质,所以上述定理也可以改写成

定理 3.2′ 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$,则 n 阶行列式 $|\boldsymbol{A}|$ 等于它的任意列的各元素与其相应的代数余子式的乘积之和.即对任意 $1\leqslant j\leqslant n$.有

$$|\boldsymbol{A}| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{i=1}^{n} a_{ij}A_{ij}.$$

由上面定理,n 阶行列式可以化成 $n-1$ 阶行列式进行计算,这个定理称为**行列式按行(列)展开定理**,这提供了计算行列式的一个有效方法,当然我们需要通过行列式的性质(即作行或列的初等变换)把某行(列)上的元素化成零较多,则按此行(列)展开便可减少计算量.

例 3.8 计算

(1) $\begin{vmatrix} 1 & 2 & 1 & 4 \\ 0 & -1 & 2 & 1 \\ 1 & 0 & 1 & 3 \\ 0 & 1 & 3 & 1 \end{vmatrix}$; (2) $\begin{vmatrix} 1 & 0 & 2 & 2 & 1 \\ -1 & 2 & 0 & 2 & 0 \\ 3 & 1 & 4 & 1 & 2 \\ 2 & -2 & 1 & 1 & 4 \\ 3 & -1 & -2 & 4 & 5 \end{vmatrix}$.

解 (1) 原式 $=\begin{vmatrix} 1 & 2 & 1 & 4 \\ 0 & -1 & 2 & 1 \\ 0 & -2 & 0 & -1 \\ 0 & 1 & 3 & 1 \end{vmatrix} = 1\cdot\begin{vmatrix} -1 & 2 & 1 \\ -2 & 0 & -1 \\ 1 & 3 & 1 \end{vmatrix} = \begin{vmatrix} -3 & 2 & 1 \\ 0 & 0 & -1 \\ -1 & 3 & 1 \end{vmatrix}$

$$=(-1)\cdot(-1)\begin{vmatrix} -3 & 2 \\ -1 & 3 \end{vmatrix} = -9+2=-7.$$

(2) 原式 $=\begin{vmatrix} 1 & 2 & 2 & 4 & 1 \\ -1 & 0 & 0 & 0 & 0 \\ 3 & 7 & 4 & 7 & 2 \\ 2 & 2 & 1 & 5 & 4 \\ 3 & 5 & -2 & 10 & 5 \end{vmatrix} = \begin{vmatrix} 2 & 2 & 4 & 1 \\ 7 & 4 & 7 & 2 \\ 2 & 1 & 5 & 4 \\ 5 & -2 & 10 & 5 \end{vmatrix}$

$$= \begin{vmatrix} -2 & 2 & -6 & -7 \\ -1 & 4 & -13 & -14 \\ 0 & 1 & 0 & 0 \\ 9 & -2 & 20 & 13 \end{vmatrix} = -\begin{vmatrix} -2 & -6 & -7 \\ -1 & -13 & -14 \\ 9 & 20 & 13 \end{vmatrix} = -\begin{vmatrix} -2 & -6 & -7 \\ 3 & -1 & 0 \\ 5 & 8 & -1 \end{vmatrix}$$

$$=-\begin{vmatrix}-20&-6&-7\\0&-1&0\\29&8&-1\end{vmatrix}=\begin{vmatrix}-20&-7\\29&-1\end{vmatrix}=20+203=223.$$

例 3.9 计算行列式

$$\begin{vmatrix}a&b&0&\cdots&0&0\\0&a&b&\cdots&0&0\\0&0&a&\cdots&0&0\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&a&b\\b&0&0&\cdots&0&a\end{vmatrix}.$$

解 把行列式按第一列展开得

$$原式=a\cdot\begin{vmatrix}a&b&\cdots&0&0\\0&a&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&a&b\\0&0&\cdots&0&a\end{vmatrix}+b\cdot(-1)^{n+1}\begin{vmatrix}b&0&\cdots&0&0\\a&b&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&b&0\\0&0&\cdots&a&b\end{vmatrix}$$

$$=a^n+(-1)^{n+1}b^n.$$

例 3.10 证明 $n(n\geqslant 2)$ 阶范德蒙德(Vandermonde)行列式

$$D_n=\begin{vmatrix}1&1&1&\cdots&1\\x_1&x_2&x_3&\cdots&x_n\\x_1^2&x_2^2&x_3^2&\cdots&x_n^2\\\vdots&\vdots&\vdots&&\vdots\\x_1^{n-1}&x_2^{n-1}&x_3^{n-1}&\cdots&x_n^{n-1}\end{vmatrix}=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

证明 对 n 作归纳.当 $n=2$ 时,

$$D_2=\begin{vmatrix}1&1\\x_1&x_2\end{vmatrix}=x_2-x_1,$$

结论正确.假设对于 $n-1$ 阶范德蒙德行列式结论正确,则对 n 阶范德蒙德行列式来说,第 n 行加上第 $n-1$ 行的 $-x_n$ 倍,第 $n-1$ 行加上第 $n-2$ 行的 $-x_n$ 倍.以此类推,第 2 行加上第 1 行的 $-x_n$ 倍,则有

$$D_n=\begin{vmatrix}1&1&\cdots&1&1\\x_1-x_n&x_2-x_n&\cdots&x_{n-1}-x_n&0\\x_1(x_1-x_n)&x_2(x_2-x_n)&\cdots&x_{n-1}(x_{n-1}-x_n)&0\\\vdots&\vdots&&\vdots&\vdots\\x_1^{n-3}(x_1-x_n)&x_2^{n-3}(x_2-x_n)&\cdots&x_{n-1}^{n-3}(x_{n-1}-x_n)&0\\x_1^{n-2}(x_1-x_n)&x_2^{n-2}(x_2-x_n)&\cdots&x_{n-1}^{n-2}(x_{n-1}-x_n)&0\end{vmatrix}$$

$$=(-1)^{1+n}\cdot(x_1-x_n)(x_2-x_n)\cdots(x_{n-1}-x_n)\cdot\begin{vmatrix}1&1&\cdots&1\\x_1&x_2&\cdots&x_{n-1}\\\vdots&\vdots&&\vdots\\x_1^{n-2}&x_2^{n-2}&\cdots&x_{n-1}^{n-2}\end{vmatrix}$$

$$=(x_n-x_1)(x_n-x_2)\cdots(x_n-x_{n-1})\cdot D_{n-1}.$$

由归纳假设 $D_{n-1}=\prod\limits_{1\leqslant i<j\leqslant n-1}(x_j-x_i)$.故

$$D_n=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

即结论对 n 时成立.由归纳法原理,结论对任意正整数 $n(n\geqslant2)$ 均成立. □

例 3.11 证明

$$\begin{vmatrix}a_{11}&\cdots&a_{1r}&0&\cdots&0\\\vdots&&\vdots&\vdots&&\vdots\\a_{r1}&\cdots&a_{rr}&0&\cdots&0\\c_{11}&\cdots&c_{1r}&b_{11}&\cdots&b_{1s}\\\vdots&&\vdots&\vdots&&\vdots\\c_{s1}&\cdots&c_{sr}&b_{s1}&\cdots&b_{ss}\end{vmatrix}=\begin{vmatrix}a_{11}&\cdots&a_{1r}\\\vdots&&\vdots\\a_{r1}&\cdots&a_{rr}\end{vmatrix}\cdot\begin{vmatrix}b_{11}&\cdots&b_{1s}\\\vdots&&\vdots\\b_{s1}&\cdots&b_{ss}\end{vmatrix}.$$

证明 设 $\boldsymbol{A}=(a_{ij})_{r\times r},\boldsymbol{B}=(b_{ij})_{s\times s},\boldsymbol{C}=(c_{ij})_{s\times r}$,则要证明的式子可简写为

$$\begin{vmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{C}&\boldsymbol{B}\end{vmatrix}=|\boldsymbol{A}|\cdot|\boldsymbol{B}|.$$

对 r 作归纳.

当 $r=1$ 时,把行列式对第一行展开,即有

$$\begin{vmatrix}a_{11}&0&\cdots&0\\c_{11}&b_{11}&\cdots&b_{1s}\\\vdots&\vdots&&\vdots\\c_{s1}&b_{s1}&\cdots&b_{ss}\end{vmatrix}=a_{11}\cdot\begin{vmatrix}b_{11}&\cdots&b_{1s}\\\vdots&&\vdots\\b_{s1}&\cdots&b_{ss}\end{vmatrix}=|\boldsymbol{A}|\cdot|\boldsymbol{B}|.$$

结论正确.假设结论对 $r-1$ 时成立,则对 r 时,仍把行列式按第一行展开,并把 $\boldsymbol{A}$ 中去掉第 1 行第 j 列的矩阵记为 $\boldsymbol{A}_j$, $\boldsymbol{C}$ 中去掉第 j 列的矩阵记作 $\boldsymbol{C}_j$, $1\leqslant j\leqslant r$.则有

$$\begin{vmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{C} & \boldsymbol{B}\end{vmatrix}=a_{11}\begin{vmatrix}\boldsymbol{A}_1 & \boldsymbol{O}\\ \boldsymbol{C}_1 & \boldsymbol{B}\end{vmatrix}+(-1)^{1+2}a_{12}\begin{vmatrix}\boldsymbol{A}_2 & \boldsymbol{O}\\ \boldsymbol{C}_2 & \boldsymbol{B}\end{vmatrix}+\cdots+(-1)^{1+r}a_{1r}\begin{vmatrix}\boldsymbol{A}_r & \boldsymbol{O}\\ \boldsymbol{C}_r & \boldsymbol{B}\end{vmatrix},$$

因为这里的 $\boldsymbol{A}_j$, $1\leqslant j\leqslant r$,都是 $r-1$ 阶矩阵,由归纳假设,有

$$\begin{vmatrix}\boldsymbol{A}_j & \boldsymbol{O}\\ \boldsymbol{C}_j & \boldsymbol{B}\end{vmatrix}=|\boldsymbol{A}_j|\cdot|\boldsymbol{B}|,$$

故

$$\begin{aligned}\begin{vmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{C} & \boldsymbol{B}\end{vmatrix}&=\sum_{j=1}^{r}a_{1j}(-1)^{1+j}|\boldsymbol{A}_j|\cdot|\boldsymbol{B}|\\&=\left(\sum_{j=1}^{r}a_{1j}(-1)^{1+j}|\boldsymbol{A}_j|\right)\cdot|\boldsymbol{B}|=|\boldsymbol{A}|\cdot|\boldsymbol{B}|.\end{aligned}$$

即结论对 r 时成立.由归纳法原理,结论对任意正整数 r 均成立. □

命题 3.2 n 阶行列式的任一行(列)的元素与另一行(列)对应元素的代数余子式的乘积之和等于 0,即当 $i\neq j$ 时

$$\sum_{k=1}^{n}a_{ik}A_{jk}=a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn}=0,$$

$$\sum_{k=1}^{n}a_{ki}A_{kj}=a_{1i}A_{1j}+a_{2i}A_{2j}+\cdots+a_{ni}A_{nj}=0.$$

证明 只证第一个式子. 注意到 A_{jk} 与原行列式中的第 j 行元素无关,将原行列式 $|\boldsymbol{A}|$ 的第 j 行元素换成第 i 行相应的元素,则得一个新的行列式

$$D_1=\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{in}\\ \vdots & \vdots & & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{in}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix},$$

显然, D_1 中有两行元素相同,故 $D_1=0$. 另一方面, D_1 的第 j 行元素的代数余子式与 $|\boldsymbol{A}|$ 的第 j 行对应元素的代数余子式相同.因此把 D_1 按第 j 行展开就有

$$D_1=a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn}=0.$$ □

§3.5 克拉默(Cramer)法则

考虑含有 n 个未知量,n 个方程的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases} \tag{3.1}$$

设它的系数矩阵为 $\boldsymbol{A}=(a_{ij})_{n\times n}$,再用 $\boldsymbol{B}_j$ 表示把 $\boldsymbol{A}$ 的第 j 列换成常数项 $b_1,b_2,\cdots,b_n$ 后所得的矩阵,即

$$\boldsymbol{B}_j = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}.$$

定理 3.3(克拉默法则) 若线性方程组(3.1)的系数矩阵 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|\neq 0$,则它有唯一解

$$x_1 = \frac{|\boldsymbol{B}_1|}{|\boldsymbol{A}|}, \quad x_2 = \frac{|\boldsymbol{B}_2|}{|\boldsymbol{A}|}, \quad \cdots, \quad x_n = \frac{|\boldsymbol{B}_n|}{|\boldsymbol{A}|}.$$

证明 先证$\left(\frac{|\boldsymbol{B}_1|}{|\boldsymbol{A}|},\frac{|\boldsymbol{B}_2|}{|\boldsymbol{A}|},\cdots,\frac{|\boldsymbol{B}_n|}{|\boldsymbol{A}|}\right)$是方程组的一个解.对 $1\leqslant j\leqslant n$,把每个 $|\boldsymbol{B}_j|$ 按第 j 列展开得

$$|\boldsymbol{B}_j| = \sum_{k=1}^{n} b_k A_{kj},$$

其中 A_{kj} 为 $\boldsymbol{A}$ 中第 k 行第 j 列元素 a_{kj} 的代数余子式,将$\left(\frac{|\boldsymbol{B}_1|}{|\boldsymbol{A}|},\frac{|\boldsymbol{B}_2|}{|\boldsymbol{A}|},\cdots,\frac{|\boldsymbol{B}_n|}{|\boldsymbol{A}|}\right)$代入方程组(3.1)中的第 i 个方程.

$$\begin{aligned} \text{左端} &= \sum_{j=1}^{n} a_{ij}\frac{|\boldsymbol{B}_j|}{|\boldsymbol{A}|} = \frac{1}{|\boldsymbol{A}|}\sum_{j=1}^{n} a_{ij}\sum_{k=1}^{n} b_k A_{kj} \\ &= \frac{1}{|\boldsymbol{A}|}\sum_{k=1}^{n} b_k \cdot \sum_{j=1}^{n} a_{ij}A_{kj}. \end{aligned}$$

由于

$$\sum_{j=1}^{n} a_{ij}A_{kj} = \begin{cases} |A|, & k=i, \\ 0, & k \neq i. \end{cases}$$

所以,左端 $=\frac{1}{|A|}\cdot b_i\cdot|A|=b_i=$ 右端,这便证出 $\left(\frac{|B_1|}{|A|},\frac{|B_2|}{|A|},\cdots,\frac{|B_n|}{|A|}\right)$ 是方程组(3.1)的一个解.

再证解的唯一性.即证明若 $(c_1,c_2,\cdots,c_n)$ 是(3.1)的解,则一定有 $c_j=\frac{|B_j|}{|A|}, j=1,2,\cdots,n$.事实上,由于

$$a_{i1}c_1+a_{i2}c_2+\cdots+a_{in}c_n=b_i,\quad 1\leqslant i\leqslant n,$$

故 $\sum_{i=1}^{n}(a_{i1}c_1+a_{i2}c_2+\cdots+a_{in}c_n)A_{ij}=\sum_{i=1}^{n}b_iA_{ij},\quad 1\leqslant j\leqslant n.$

即

$$c_1\cdot\sum_{i=1}^{n}a_{i1}A_{ij}+c_2\cdot\sum_{i=1}^{n}a_{i2}A_{ij}+\cdots+c_n\cdot\sum_{i=1}^{n}a_{in}A_{ij}=|B_j|.$$

由于

$$\sum_{i=1}^{n} a_{ik}A_{ij} = \begin{cases} |A|, & k=j, \\ 0, & k \neq j. \end{cases}$$

所以

$$c_j\cdot|A|=|B_j|,$$

即

$$c_j=\frac{|B_j|}{|A|},\quad 1\leqslant j\leqslant n.$$ □

注意,因为当 n 很大时计算 n 阶行列式是不容易的,所以克拉默法则只是具有理论上的意义,具体解线性方程组还是用第二章中讲过的高斯消元法.

若 $b_1=b_2=\cdots b_n=0$,即方程组(3.1)为齐次线性方程组,这时 $|B_j|=0, j=1,2,\cdots,n$.所以若其系数矩阵的行列式不等于0,则它只有零解,或者写成下面的定理.

定理 3.4 若一个有 n 个未知量、n 个方程的齐次线性方程组有非零解,则其系数矩阵的行列式一定为0.

下面考虑它的逆命题,即一个有 n 个未知量、n 个方程的齐次线性方程组的系数矩阵的行列式为0,则它是否有非零解?

从第二章中的结论知道,将这个方程组的系数矩阵做初等行变换化为阶梯形.设这个阶梯形矩阵中的非零行个数为 r,则方程组有非零解当且仅当 $r<n$.

对矩阵 A 做初等行变换得到的矩阵的行列式是 $c\cdot|A|$,其中 $c\neq0$,这可从行列式的性质容易看出.又显然阶梯形方阵 J 一定是上三角矩阵,它的行列式为其主对角线上元素的乘积,又 J 的非零行个数小于 n 等价于其主对角线上有元素0,即等价于 $|J|=0$.由于 $|J|=c|A|, c\neq0$,所以这也等价于 $|A|=0$.

所以,若一个含 n 个未知量、n 个方程的齐次线性方程组的系数矩阵的行列式为0,则把

它的系数矩阵做初等行变换后得到的阶梯形矩阵的行列式也为 0,所以这个阶梯形矩阵的非零行个数小于 n.即这个齐次线性方程组有非零解.这样,我们就证明了下述结论.

定理 3.5 一个有 n 个未知量、n 个方程的齐次线性方程组有非零解的充要条件是它的系数矩阵的行列式为 0,或者它只有零解的充要条件是它的系数矩阵的行列式不为 0.

习 题 三

1. 求下列 n 元排列的逆序数:

(1) $(n-1)(n-2)\cdots 21n$; (2) $23\cdots(n-1)n1$.

2. 如果 n 元排列 $j_1j_2\cdots j_{n-1}j_n$ 的逆序数为 r,求 n 元排列 $j_nj_{n-1}\cdots j_2j_1$ 的逆序数.

3. 计算下列行列式:

(1) $\begin{vmatrix} 1 & 4 & 2 \\ 3 & 5 & 1 \\ 2 & 1 & 6 \end{vmatrix}$; (2) $\begin{vmatrix} 2 & -1 & 5 \\ 3 & 1 & -2 \\ 1 & 4 & 6 \end{vmatrix}$;

(3) $\begin{vmatrix} 1 & 0 & -3 & 2 \\ -4 & -1 & 0 & -5 \\ 2 & 3 & -1 & -6 \\ 3 & 3 & -4 & 1 \end{vmatrix}$; (4) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$.

4. 计算下列 n 阶行列式:

(1) $\begin{vmatrix} a & 1 & 1 & \cdots & 1 \\ 1 & a & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & a \end{vmatrix}$; (2) $\begin{vmatrix} a_1+b_1 & a_1+b_2 & \cdots & a_1+b_n \\ a_2+b_1 & a_2+b_2 & \cdots & a_2+b_n \\ \vdots & \vdots & & \vdots \\ a_n+b_1 & a_n+b_2 & \cdots & a_n+b_n \end{vmatrix}$.

5. 求证:

$$\begin{vmatrix} x_1+y_1 & y_1+z_1 & z_1+x_1 \\ x_2+y_2 & y_2+z_2 & z_2+x_2 \\ x_3+y_3 & y_3+z_3 & z_3+x_3 \end{vmatrix} = 2\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

6. 求证:

$$\begin{vmatrix} x & x^2 & yz \\ y & y^2 & xz \\ z & z^2 & xy \end{vmatrix} = (xy+yz+zx)(x-y)(y-z)(z-x).$$

7. 已知行列式$\begin{vmatrix} 2 & 1 & -5 & 1 \\ 1 & -3 & 0 & -6 \\ 0 & 2 & -1 & 2 \\ 1 & 4 & -7 & 6 \end{vmatrix}$,设 A_{ij}是 a_{ij}的代数余子式.求

(1) $A_{41}+4A_{42}-7A_{43}+6A_{44}$;

(2) $A_{11}-3A_{12}-6A_{14}$;

(3) $A_{41}+A_{42}+A_{43}+A_{44}$.

8. 已知平面上三条不同直线的方程分别为$\begin{cases} l_1: ax+2by+3c=0, \\ l_2: bx+2cy+3a=0, \\ l_3: cx+2ay+3b=0, \end{cases}$求证:这三条直线交于一点的充要条件是 $a+b+c=0$.

9. 当 a,b 取什么值时,下述线性方程组有唯一解?

$$\begin{cases} ax_1 + x_2 + x_3 = 2, \\ x_1 + bx_2 + x_3 = 1, \\ x_1 + 2bx_2 + x_3 = 2. \end{cases}$$

10. 讨论下述线性方程组何时有唯一解?有无穷多个解?无解?

$$\begin{cases} ax_1 + x_2 + x_3 = 2, \\ x_1 + bx_2 + x_3 = 1, \\ x_1 + 2bx_2 + x_3 = 1. \end{cases}$$

第四章 矩阵代数

矩阵在很多实际问题中有重要应用.第二章中我们已经用之来解线性方程组,主要操作就是矩阵的初等行变换.本章将介绍矩阵的运算以及运算性质等.

§4.1 矩阵及其运算

4.1.1 加法和数乘

定义 4.1 设矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$,$\boldsymbol{B}=(b_{ij})_{s\times t}$,则当 $m=s$,$n=t$ 且对任意 $1\leqslant i\leqslant m$,$1\leqslant j\leqslant n$ 都有 $a_{ij}=b_{ij}$时称**矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相等**.记作 $\boldsymbol{A}=\boldsymbol{B}$.

所以只有当两个矩阵行数相同,列数相同且对应位置元素也相等,即它们写出来是完全一样时才称这两个矩阵相等.所以对矩阵做初等行变换得到的矩阵与原矩阵是不相等的.

定义 4.2 设矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$,$\boldsymbol{B}=(b_{ij})_{m\times n}$,称矩阵 $\boldsymbol{C}=(c_{ij})_{m\times n}$,其中 $c_{ij}=a_{ij}+b_{ij}$,为**矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和**.记作 $\boldsymbol{C}=\boldsymbol{A}+\boldsymbol{B}$.

只有行数相同、列数也相同的两个矩阵可以相加.和矩阵中的元素是这两个矩阵对应位置元素的和.

元素全为 0 的矩阵称为**零矩阵**,记作 $\boldsymbol{O}$.对于矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$,定义 $-\boldsymbol{A}=(-a_{ij})_{m\times n}$,称它为 $\boldsymbol{A}$ 的**负矩阵**.即把 $\boldsymbol{A}$ 中每个元素都变为它的相反数.

容易验证矩阵加法满足如下运算法则,其中 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 为任意 $m\times n$ 矩阵.

(1) 交换律 即 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$.

(2) 结合律 即 $\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})=(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}$.

(3) $\boldsymbol{A}+\boldsymbol{O}=\boldsymbol{A}$.

(4) $\boldsymbol{A}+(-\boldsymbol{A})=\boldsymbol{O}$.

有了负矩阵.可以定义**矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的差**为 $\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{A}+(-\boldsymbol{B})$,即对应位置元素相减所得到的矩阵.

例 4.1 设矩阵 $\boldsymbol{A}=\begin{pmatrix}2 & a\\0 & 5\end{pmatrix}$,$\boldsymbol{B}=\begin{pmatrix}b & -1\\c & 7\end{pmatrix}$且 $\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}3 & 4\\5 & 12\end{pmatrix}$.求 a,b,c.

解 由于 $\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}2+b & a-1\\c & 12\end{pmatrix}=\begin{pmatrix}3 & 4\\5 & 12\end{pmatrix}$.

所以 $2+b=3,a-1=4,c=5.$

解得 $a=5,b=1,c=5.$

定义 4.3 设有数 k 和矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$,则称矩阵$(ka_{ij})_{m\times n}$为**数 k 与矩阵 $\boldsymbol{A}$ 的乘积**,记作 $k\boldsymbol{A}$.

数与矩阵的乘积也简称为**数乘**,即用数 k 去乘矩阵中的每个元素.

容易验证矩阵的数乘满足如下运算法则,其中 k,l 为任意数,$\boldsymbol{A}$ 为任意矩阵.

(5) $1\cdot\boldsymbol{A}=\boldsymbol{A}$.

(6) $(kl)\boldsymbol{A}=k(l\boldsymbol{A})$.

矩阵的加法和数乘满足下面法则,其中 k,l 为任意数,$\boldsymbol{A},\boldsymbol{B}$ 为任意 $m\times n$ 矩阵.

(7) $(k+l)\boldsymbol{A}=k\boldsymbol{A}+l\boldsymbol{A}$.

(8) $k(\boldsymbol{A}+\boldsymbol{B})=k\boldsymbol{A}+k\boldsymbol{B}$.

4.1.2 乘法

定义 4.4 设矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$,$\boldsymbol{B}=(b_{ij})_{n\times s}$,称矩阵 $\boldsymbol{D}=(d_{ij})_{m\times s}$,其中,

$$d_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}=\sum_{k=1}^{n}a_{ik}b_{kj},1\leqslant i\leqslant m,1\leqslant j\leqslant s,$$

为**矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积**,记作 $\boldsymbol{D}=\boldsymbol{AB}$.

由定义可以看出.只有矩阵 $\boldsymbol{A}$ 的列数等于矩阵 $\boldsymbol{B}$ 的行数时,$\boldsymbol{AB}$ 才有意义且 $\boldsymbol{AB}$ 中第 i 行 j 列的元素为矩阵 $\boldsymbol{A}$ 的第 i 行与矩阵 $\boldsymbol{B}$ 的第 j 列对应元素乘积的和.

例 4.2 设 $\boldsymbol{A}=(a_1,a_2,\cdots,a_n)$,$\boldsymbol{B}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$.计算 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$.

解 $\boldsymbol{AB}=(a_1b_1+a_2b_2+\cdots+a_nb_n)_{1\times 1}$

$$BA=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}(a_1,a_2,\cdots,a_n)=\begin{pmatrix}b_1a_1&b_1a_2&\cdots&b_1a_n\\b_2a_1&b_2a_2&\cdots&b_2a_n\\\vdots&\vdots&&\vdots\\b_na_1&b_na_2&\cdots&b_na_n\end{pmatrix}.$$

此例中 $\boldsymbol{AB}$ 为 1×1 矩阵，而 $\boldsymbol{BA}$ 为 $n\times n$ 矩阵，当 $n\geqslant2$ 时，自然这两个矩阵不相等.

例 4.3 设 $\boldsymbol{A}=\begin{pmatrix}1&1\\-1&-1\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}1&-1\\-1&1\end{pmatrix}$.求 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$.

解
$$\boldsymbol{AB}=\begin{pmatrix}1&1\\-1&-1\end{pmatrix}\begin{pmatrix}1&-1\\-1&1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}=\boldsymbol{O}.$$

$$\boldsymbol{BA}=\begin{pmatrix}1&-1\\-1&1\end{pmatrix}\begin{pmatrix}1&1\\-1&-1\end{pmatrix}=\begin{pmatrix}2&2\\-2&-2\end{pmatrix}\neq\boldsymbol{O},$$

由上面例子可以看出，矩阵乘积没有交换性，即 $\boldsymbol{AB}$ 不一定等于 $\boldsymbol{BA}$.若 $\boldsymbol{AB}=\boldsymbol{BA}$，则称 $\boldsymbol{A},\boldsymbol{B}$ 可交换.同时也存在 $\boldsymbol{A}\neq\boldsymbol{O},\boldsymbol{B}\neq\boldsymbol{O}$.但是 $\boldsymbol{AB}=\boldsymbol{O}$，所以由 $\boldsymbol{AB}=\boldsymbol{O}$ 不能得出 $\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{B}=\boldsymbol{O}$.

下面来看矩阵乘法所满足的运算法则.

命题 4.1 矩阵乘积满足结合律.即当乘积可以进行时，对任意 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 有

$$\boldsymbol{A}(\boldsymbol{BC})=(\boldsymbol{AB})\boldsymbol{C}.$$

证明 设 $\boldsymbol{A}=(a_{ij})_{m\times n}$，$\boldsymbol{B}=(b_{ij})_{n\times s}$，$\boldsymbol{C}=(c_{ij})_{s\times t}$.并记 $\boldsymbol{AB}=\boldsymbol{P}=(p_{ij})$，$\boldsymbol{BC}=\boldsymbol{Q}=(q_{ij})$，则 $\boldsymbol{P}$ 为 $m\times s$ 矩阵，而 $\boldsymbol{Q}$ 为 $n\times t$ 矩阵.所以 $\boldsymbol{A}(\boldsymbol{BC})=\boldsymbol{AQ}$ 为 $m\times t$ 矩阵.$(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{PC}$ 也是 $m\times t$ 矩阵.

矩阵 $\boldsymbol{A}(\boldsymbol{BC})=\boldsymbol{AQ}$ 的第 i 行第 j 列位置元素为

$$d_{ij}=\sum_{k=1}^{n}a_{ik}q_{kj}=\sum_{k=1}^{n}a_{ik}\left(\sum_{l=1}^{s}b_{kl}c_{lj}\right)=\sum_{k=1}^{n}\sum_{l=1}^{s}a_{ik}b_{kl}c_{lj}.$$

而矩阵 $(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{PC}$ 的第 (i,j) 元素为

$$e_{ij}=\sum_{l=1}^{s}p_{il}c_{lj}=\sum_{l=1}^{s}\left(\sum_{k=1}^{n}a_{ik}b_{kl}\right)c_{lj}=\sum_{l=1}^{s}\sum_{k=1}^{n}a_{ik}b_{kl}c_{lj}=d_{ij}.$$

即它们对应位置的元素均相等.所以 $\boldsymbol{A}(\boldsymbol{BC})=(\boldsymbol{AB})\boldsymbol{C}$. □

命题 4.2 矩阵的加法和乘法满足分配律，即对矩阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 有

$$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC},\quad(\boldsymbol{B}+\boldsymbol{C})\boldsymbol{A}=\boldsymbol{BA}+\boldsymbol{CA}.$$

证明 根据定义推导即可. □

定义 4.5 主对角线外元素均为 0 的 $n\times n$ 矩阵称为**对角矩阵**.主对角线上元素分别为 $d_1,d_2,\cdots,d_n$ 的对角矩阵简记为 $\mathrm{diag}(d_1,d_2,\cdots,d_n)$.主对角线上元素都是 1 的对角矩阵称为**单位矩阵**，记为 $\boldsymbol{I}_n$(或当上下文中清楚矩阵的阶时记为 $\boldsymbol{I}$).主对角线上元素都相等的对角矩阵称为**数量矩阵**，即矩阵 $\mathrm{diag}(d,d,\cdots,d)$，也常记为 $d\boldsymbol{I}_n$(或 $d\boldsymbol{I}$).

容易验证

$$\operatorname{diag}(d_1,d_2,\cdots,d_m)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} d_1a_{11} & d_1a_{12} & \cdots & d_1a_{1n} \\ d_2a_{21} & d_2a_{22} & \cdots & d_2a_{2n} \\ \vdots & \vdots & & \vdots \\ d_ma_{m1} & d_ma_{m2} & \cdots & d_ma_{mn} \end{pmatrix}.$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\operatorname{diag}(d_1,d_2,\cdots,d_n) = \begin{pmatrix} d_1a_{11} & d_2a_{12} & \cdots & d_na_{1n} \\ d_1a_{21} & d_2a_{22} & \cdots & d_na_{2n} \\ \vdots & \vdots & & \vdots \\ d_1a_{m1} & d_2a_{m2} & \cdots & d_na_{mn} \end{pmatrix}.$$

即一个对角矩阵左乘任一矩阵 $\boldsymbol{A}$ 相当于用对角矩阵的主对角线上的元素依次乘以 $\boldsymbol{A}$ 的对应的行,而一个对角矩阵右乘任一矩阵 $\boldsymbol{A}$ 相当于用对角矩阵的主对角线上的元素依次乘以 $\boldsymbol{A}$ 的对应的列.特别地,对任一 $m\times n$ 矩阵 $\boldsymbol{A}$,有

(1) $\boldsymbol{I}_m\boldsymbol{A}=\boldsymbol{A}\boldsymbol{I}_n=\boldsymbol{A}$,

(2) $(d\boldsymbol{I}_m)\boldsymbol{A}=\boldsymbol{A}(d\boldsymbol{I}_n)=d\boldsymbol{A}$.

再利用矩阵乘积的结合律可得,对任意 $m\times n$ 矩阵 $\boldsymbol{A}$ 和 $n\times s$ 矩阵 $\boldsymbol{B}$,有

$$(k\boldsymbol{A})\boldsymbol{B}=(\boldsymbol{A}(k\boldsymbol{I}_n))\boldsymbol{B}=\boldsymbol{A}((k\boldsymbol{I}_n)\boldsymbol{B})=\boldsymbol{A}(k\boldsymbol{B}),$$

以及

$$(k\boldsymbol{A})\boldsymbol{B}=((k\boldsymbol{I}_m)\boldsymbol{A})\boldsymbol{B}=(k\boldsymbol{I}_m)(\boldsymbol{A}\boldsymbol{B})=k(\boldsymbol{A}\boldsymbol{B}).$$

这便证明了下面的结论.

命题 4.3 矩阵的数乘和乘积满足

$$k(\boldsymbol{A}\boldsymbol{B})=(k\boldsymbol{A})\boldsymbol{B}=\boldsymbol{A}(k\boldsymbol{B}).$$

因为矩阵的乘法满足结合律.所以 $(\boldsymbol{A}\boldsymbol{B})\boldsymbol{C}$ 可以记成 $\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}$,当 $\boldsymbol{A}$ 是方阵时.对于正整数 n,定义 $\boldsymbol{A}^n=\underbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}}_{n个}$,即 n 个 $\boldsymbol{A}$ 的连乘.容易验证,对任意方阵 $\boldsymbol{A}$ 和正整数 k 和 l,有

$$\boldsymbol{A}^k\boldsymbol{A}^l=\boldsymbol{A}^{k+l},\quad (\boldsymbol{A}^k)^l=\boldsymbol{A}^{kl}.$$

对于任意多项式 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 和方阵 $\boldsymbol{A}$,定义 $f(\boldsymbol{A})=a_n\boldsymbol{A}^n+a_{n-1}\boldsymbol{A}^{n-1}+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{I}$,并称其为**矩阵 $\boldsymbol{A}$ 的多项式**.

4.1.3 转置

定义 4.6 设 $\boldsymbol{A}=(a_{ij})_{m\times n}$,定义 $\boldsymbol{A}$ 的转置为一个 $n\times m$ 矩阵,其第 i 行第 j 列元素等于 $\boldsymbol{A}$ 的第 j 行第 i 列元素 a_{ji},记为 $\boldsymbol{A}^{\mathrm{T}}$(或 $\boldsymbol{A}'$).即 $\boldsymbol{A}^{\mathrm{T}}$ 是把 $\boldsymbol{A}$ 的行、列互换所得到的矩阵.

命题 4.4 对于任意数 k 及矩阵 $\boldsymbol{A},\boldsymbol{B}$,当相关运算有定义时,矩阵的转置具有下列性质:

(1) $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{A}$.

(2) $(k\boldsymbol{A})^{\mathrm{T}}=k\boldsymbol{A}^{\mathrm{T}}$.

(3) $(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}$.

(4) $(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

证明 (1),(2),(3)是显然的.下面证明(4).

设 $\boldsymbol{A}=(a_{ij})_{m\times n}$,$\boldsymbol{B}=(b_{ij})_{n\times s}$,则$(\boldsymbol{AB})^{\mathrm{T}}$与$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$都是 $s\times m$ 矩阵.又$(\boldsymbol{AB})^{\mathrm{T}}$的第 i 行第 j 列元素为 $\boldsymbol{AB}$ 的第 j 行第 i 列元素,即为

$$a_{j1}b_{1i}+a_{j2}b_{2i}+\cdots+a_{jn}b_{ni}.$$

而 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ 的第 i 行第 j 列元素为 $\boldsymbol{B}^{\mathrm{T}}$ 的第 i 行元素,即 $\boldsymbol{B}$ 的第 i 列元素 $b_{1i},b_{2i},\cdots,b_{ni}$与 $\boldsymbol{A}^{\mathrm{T}}$ 的第 j 列元素,即 $\boldsymbol{A}$ 的第 j 行元素 $a_{j1},a_{j2},\cdots,a_{jn}$对应项乘积之和,即

$$b_{1i}a_{j1}+b_{2i}a_{j2}+\cdots+b_{ni}a_{jn}.$$

因此对任意 $1\leqslant i\leqslant s,1\leqslant j\leqslant m$,$(\boldsymbol{AB})^{\mathrm{T}}$的第 i 行第 j 列元素与 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ 的第 i 行第 j 列元素相等,故有

$$(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}.$$ □

若矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为**对称矩阵**.若满足 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为**反对称矩阵**.

4.1.4 矩阵的运算与行列式

下面设出现的矩阵为方阵,自然这样的矩阵有行列式.由行列式性质,显然有 $|\boldsymbol{A}^{\mathrm{T}}|=|\boldsymbol{A}|$,$|k\boldsymbol{A}|=k^n|\boldsymbol{A}|$.所以奇数阶反对称矩阵的行列式一定为0.

命题 4.5 设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶方阵,则有 $|\boldsymbol{AB}|=|\boldsymbol{A}|\cdot|\boldsymbol{B}|$.

证明 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$,$\boldsymbol{B}=(b_{ij})_{n\times n}$,$\boldsymbol{AB}=(c_{ij})_{n\times n}$,计算下面这个 $2n$ 阶行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}.$$

由例 3.11 知 $D=|\boldsymbol{A}|\cdot|\boldsymbol{B}|$.另一方面,把第 $n+1$ 行的 a_{11}倍,第 $n+2$ 行的 a_{12}倍,$\cdots$,第 $2n$ 行的 a_{1n}倍加到第 1 行,得

$$D=\begin{vmatrix} 0 & 0 & \cdots & 0 & c_{11} & c_{12} & \cdots & c_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}.$$

再把第 $n+1$ 行的 a_{21} 倍，第 $n+2$ 行的 a_{22} 倍，…，第 $2n$ 行的 a_{2n} 倍加到第 2 行. 并以此类推，一直把第 $n+1$ 行的第 a_{n1} 倍，第 $n+2$ 行的第 a_{n2} 倍，…，第 $2n$ 行的 a_{nn} 倍加到第 n 行. 最后得

$$D=\begin{vmatrix} 0 & 0 & \cdots & 0 & c_{11} & c_{12} & \cdots & c_{1n} \\ 0 & 0 & \cdots & 0 & c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & c_{n1} & c_{n2} & \cdots & c_{nn} \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}.$$

将第 $n+1$ 列与它前面的 n 列依次交换把第它换到第 1 列，再把第 $n+2$ 列与它前面的 n 列依次交换把它换到第 2 列. 以此类推，把第 $2n$ 列与它前面的 n 列依次交换把它换到第 n 列，得

$$D=\underbrace{(-1)^n\cdot(-1)^n\cdot\cdots\cdot(-1)^n}_{n个}\begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} & 0 & 0 & \cdots & 0 \\ c_{21} & c_{22} & \cdots & c_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} & 0 & 0 & \cdots & 0 \\ b_{11} & b_{12} & \cdots & b_{1n} & -1 & 0 & \cdots & 0 \\ b_{21} & b_{22} & \cdots & b_{2n} & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} & 0 & 0 & \cdots & -1 \end{vmatrix}.$$

再利用例 3.11 得

$$D=(-1)^{n^2}\cdot\begin{vmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\vdots&\vdots&&\vdots\\c_{n1}&c_{n2}&\cdots&c_{nn}\end{vmatrix}\cdot\begin{vmatrix}-1&0&\cdots&0\\0&-1&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&-1\end{vmatrix}=(-1)^{n^2}\cdot(-1)^n\cdot|\boldsymbol{AB}|$$

$$=(-1)^{n^2+n}|\boldsymbol{AB}|=|\boldsymbol{AB}|.$$

所以

$$|\boldsymbol{AB}|=|\boldsymbol{A}|\cdot|\boldsymbol{B}|.$$ □

注意 一般说来，$|\boldsymbol{A}+\boldsymbol{B}|\neq|\boldsymbol{A}|+|\boldsymbol{B}|$.

§4.2 分块矩阵

4.2.1 分块矩阵

只有一行的矩阵称为行向量，只有一列的矩阵称为列向量.设 $m\times n$ 矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$，记 $\boldsymbol{\alpha}_i=(a_{i1},a_{i2},\cdots,a_{in})$，$1\leqslant i\leqslant m$，称其为 $\boldsymbol{A}$ 的第 i 个行向量，而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 称为 $\boldsymbol{A}$ 的行向量组.类似地，记

$$\boldsymbol{\beta}_j=\begin{pmatrix}a_{1j}\\a_{2j}\\\vdots\\a_{mj}\end{pmatrix},\quad 1\leqslant j\leqslant n,$$

称其为 $\boldsymbol{A}$ 的第 j 个列向量.而 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 称为 $\boldsymbol{A}$ 的列向量组.这时矩阵 $\boldsymbol{A}$ 可记为

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\vdots\\\boldsymbol{\alpha}_m\end{pmatrix}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n).$$

设 $\boldsymbol{B}=(b_{ij})_{n\times s}$，则 $\boldsymbol{AB}$ 的第 j 个 $(1\leqslant j\leqslant s)$ 列向量为

$$\begin{pmatrix} a_{11}b_{1j}+a_{12}b_{2j}+\cdots+a_{1n}b_{nj} \\ a_{21}b_{1j}+a_{22}b_{2j}+\cdots+a_{2n}b_{nj} \\ \vdots \\ a_{m1}b_{1j}+a_{m2}b_{2j}+\cdots+a_{mn}b_{nj} \end{pmatrix} = b_{1j}\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + b_{2j}\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + b_{nj}\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

$$= b_{1j}\boldsymbol{\beta}_1 + b_{2j}\boldsymbol{\beta}_2 + \cdots + b_{nj}\boldsymbol{\beta}_n,$$

其中 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 为矩阵 $\boldsymbol{A}$ 的列向量组,上面这个结论可写成如下形式

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1s} \\ b_{21} & b_{22} & \cdots & b_{2s} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{ns} \end{pmatrix}$$

$$=(b_{11}\boldsymbol{\beta}_1+b_{21}\boldsymbol{\beta}_2+\cdots+b_{n1}\boldsymbol{\beta}_n, b_{12}\boldsymbol{\beta}_1+b_{22}\boldsymbol{\beta}_2+\cdots+b_{n2}\boldsymbol{\beta}_n,\cdots,b_{1s}\boldsymbol{\beta}_1+b_{2s}\boldsymbol{\beta}_2+\cdots+b_{ns}\boldsymbol{\beta}_n).$$

所以,虽然这时 $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$ 还不是矩阵(因为它里面每个位置的元素都是向量,不是数).但形式上看,上面这个式子与普通矩阵乘法的定义是完全一致的.类似地,利用矩阵的转置,可以证明

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} \boldsymbol{\gamma}_1 \\ \boldsymbol{\gamma}_2 \\ \vdots \\ \boldsymbol{\gamma}_n \end{pmatrix} = \begin{pmatrix} a_{11}\boldsymbol{\gamma}_1+a_{12}\boldsymbol{\gamma}_2+\cdots+a_{1n}\boldsymbol{\gamma}_n \\ a_{21}\boldsymbol{\gamma}_1+a_{22}\boldsymbol{\gamma}_2+\cdots+a_{2n}\boldsymbol{\gamma}_n \\ \vdots \\ a_{m1}\boldsymbol{\gamma}_1+a_{m2}\boldsymbol{\gamma}_2+\cdots+a_{mn}\boldsymbol{\gamma}_n \end{pmatrix},$$

这里 $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_n$ 是乘积右边矩阵的行向量组.这与普通矩阵乘法的定义在形式上也是一样的.

一般地,把行、列数较大的矩阵分块处理,可以简化矩阵的运算(特别是乘法运算).对一个 $m\times n$ 矩阵 $\boldsymbol{A}$,将它的 m 行分成 p 个部分(每一部分称为一个行条),将它的 n 列分成 q 个部分(每一个部分称为一个列条),这些行条和列条将 $\boldsymbol{A}$ 分成 pq 个部分,即

$$\boldsymbol{A}=\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1q} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2q} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{p1} & \boldsymbol{A}_{p2} & \cdots & \boldsymbol{A}_{pq} \end{pmatrix}, \tag{4.1}$$

其中的每个部分 $\boldsymbol{A}_{ij}(1\leqslant i\leqslant p,1\leqslant j\leqslant q)$ 都是一个矩阵.从形式上看,$\boldsymbol{A}$ 就是以这些小的矩阵为元素的 $p\times q$ 矩阵,这样的矩阵称为**分块矩阵**.

分块矩阵的加法和数乘就是对应块相加,以及每个小块都同时用同一个数进行数乘.又若 $\boldsymbol{A}$ 为形如(4.1)的分块矩阵,则

$$A^{\mathrm{T}}=\begin{pmatrix} A_{11}^{\mathrm{T}} & A_{21}^{\mathrm{T}} & \cdots & A_{p1}^{\mathrm{T}} \\ A_{12}^{\mathrm{T}} & A_{22}^{\mathrm{T}} & \cdots & A_{p2}^{\mathrm{T}} \\ \vdots & \vdots & & \vdots \\ A_{1q}^{\mathrm{T}} & A_{2q}^{\mathrm{T}} & \cdots & A_{pq}^{\mathrm{T}} \end{pmatrix},$$

即分块矩阵转置时,先行、列互换,还要把每个小块进行转置.在进行分块矩阵乘积时,还要把矩阵的小块分得适合运算才行.设矩阵 A 形如(4.1),其中 A_{ij} 为 m_i 行 n_j 列矩阵,$1\leqslant i\leqslant p$,$1\leqslant j\leqslant q$,且 $m_1+m_2+\cdots+m_p=m$,$n_1+n_2+\cdots+n_q=n$.对于 $n\times s$ 矩阵 B,它的 n 行也要分成 q 个行条,其第 j 个行条包含 n_j 行.设 B 的 s 列分成 t 个列条,第 k 个列条包含 s_k 列.即 B 形如

$$B=\begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1t} \\ B_{21} & B_{22} & \cdots & B_{2t} \\ \vdots & \vdots & & \vdots \\ B_{q1} & B_{q2} & \cdots & B_{qt} \end{pmatrix},$$

其中 B_{ij} 为 $n_i\times s_j$ 矩阵.则 AB 可以把 A 和 B 都看成普通矩阵形式的乘法,即 AB 是一个有 p 个块行,t 个块列的分块矩阵 $C=(C_{ij})$,它的第(i,j)位置的小块为

$$C_{ij}=A_{i1}B_{1j}+A_{i2}B_{2j}+\cdots+A_{iq}B_{qj},$$

$1\leqslant i\leqslant p$,$1\leqslant j\leqslant t$.所以分块矩阵的乘法与普通矩阵乘法的定义在形式上是一样的.只是这里要注意小块 A_{ik} 和 B_{kj} 的次序,A_{ik} 必须在 B_{kj} 的左侧.

例 4.4 设

$$A=\begin{pmatrix} 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

求 A^2.

解 利用分块矩阵,把 A 写成

$$A=\begin{pmatrix} K & L \\ O & I_2 \end{pmatrix},$$

其中 $K=\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$,$L=\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.$O$ 和 I_2 分别为 2 阶零矩阵和单位矩阵.所以

$$A^2=\begin{pmatrix} K & L \\ O & I_2 \end{pmatrix}\begin{pmatrix} K & L \\ O & I_2 \end{pmatrix}=\begin{pmatrix} K^2 & KL+L \\ O & I_2 \end{pmatrix}.$$

又 $K^2=O$,$KL=O$,所以

$$A^2 = \begin{pmatrix} O & L \\ O & I_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

设线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

的系数矩阵为 $A=(a_{ij})_{m\times n}$，令 $X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$，$\beta=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$，利用矩阵乘法此方程组可写成矩阵形式

$$AX=\beta.$$

例 4.5 设 $m\times n$ 矩阵 A 和 $n\times s$ 矩阵 B 满足 $AB=O$，矩阵 B 的列向量组为 $\beta_1,\beta_2,\cdots,\beta_s$，则 $AB=O$ 当且仅当 $\beta_1,\beta_2,\cdots,\beta_s$ 都是齐次线性方程组 $AX=O$ 的解.

证明 由于 $AB=A(\beta_1,\beta_2,\cdots,\beta_s)=(A\beta_1,A\beta_2,\cdots,A\beta_s)$，所以 $AB=O$ 当且仅当 $A\beta_1=O,A\beta_2=O,\cdots,A\beta_s=O$，也即当且仅当 $\beta_1,\beta_2,\cdots,\beta_s$ 都是齐次线性方程组 $AX=O$ 的解. □

4.2.2 初等矩阵与初等变换

定义 4.7 单位矩阵 I 做一次初等变换后得到的矩阵称为**初等矩阵**.

先考虑三种初等行变换.交换单位矩阵 I 的第 i 行和第 j 行得到的矩阵记为 $P(i,j)$，即

$$P(i,j)=\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix},$$

单位矩阵 $\boldsymbol{I}$ 的第 i 行乘以非 0 数 c 得到的矩阵记为 $\boldsymbol{P}(i(c))$，即

$$\boldsymbol{P}(i(c))=\begin{pmatrix}1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1\end{pmatrix} i,$$

将单位矩阵 $\boldsymbol{I}$ 的第 i 行的 k 倍加到第 j 行记为 $\boldsymbol{P}(j,i(k))$，即

$$\boldsymbol{P}(j,i(k))=\begin{pmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \vdots & \ddots & & \\ & & k & \cdots & 1 & \\ & & & & & \ddots & \\ & & & & & & 1\end{pmatrix}\begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix}.$$

对于初等列变换来说，交换 $\boldsymbol{I}$ 的第 i 列和第 j 列得到的初等矩阵仍为 $\boldsymbol{P}(i,j)$，$\boldsymbol{I}$ 的第 i 列乘以非 0 数 c 得到的还是 $\boldsymbol{P}(i(c))$，而初等矩阵 $\boldsymbol{P}(j,i(k))$ 是将 $\boldsymbol{I}$ 的第 j 列的 k 倍加到第 i 列所得到的.从而初等矩阵也只有以上三类.

设 $m\times n$ 矩阵 $\boldsymbol{A}$ 的行向量组为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$，列向量组为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$，则在 $\boldsymbol{A}$ 的左侧乘以 m 阶初等矩阵 $\boldsymbol{P}(j,i(k))$ 可得

$$\boldsymbol{P}(j,i(k))\boldsymbol{A}=\begin{pmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \vdots & \ddots & & \\ & & k & \cdots & 1 & \\ & & & & & \ddots & \\ & & & & & & 1\end{pmatrix}\begin{pmatrix}\boldsymbol{\alpha}_1\\ \vdots\\ \boldsymbol{\alpha}_i\\ \vdots\\ \boldsymbol{\alpha}_j\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix}=\begin{pmatrix}\boldsymbol{\alpha}_1\\ \vdots\\ \boldsymbol{\alpha}_i\\ \vdots\\ \boldsymbol{\alpha}_j+k\boldsymbol{\alpha}_i\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix}.$$

在 $\boldsymbol{A}$ 的右侧乘以 n 阶初等矩阵 $\boldsymbol{P}(j,i(k))$ 可得

$$AP(j,i(k))=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_i,\cdots,\boldsymbol{\beta}_j,\cdots,\boldsymbol{\beta}_n)\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&\vdots&\ddots&&\\&&k&\cdots&1&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}$$

$$=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_i+k\boldsymbol{\beta}_j,\cdots,\boldsymbol{\beta}_j,\cdots,\boldsymbol{\beta}_n).$$

这表明,用 $\boldsymbol{P}(j,i(k))$ 左乘 $\boldsymbol{A}$,相当于把 $\boldsymbol{A}$ 的第 i 行的 k 倍加到第 j 行上,其余行不变.而用 $\boldsymbol{P}(j,i(k))$ 右乘 $\boldsymbol{A}$,相当于把 $\boldsymbol{A}$ 的第 j 列的 k 倍加到第 i 列上,其余列不变.类似地可以证明.用 $\boldsymbol{P}(i,j)$ 左(右)乘 $\boldsymbol{A}$,相当于把 $\boldsymbol{A}$ 的第 i 行(列)与第 j 行(列)互换,其余行(列)不变.用 $\boldsymbol{P}(i(c))(c\neq 0)$ 左(右)乘 $\boldsymbol{A}$,相当于用 c 乘以 $\boldsymbol{A}$ 的第 i 行(列),其余行(列)不变.把这个结论写成一个定理如下.

定理 4.1 用初等矩阵左(右)乘一个矩阵 $\boldsymbol{A}$,就相当于对 $\boldsymbol{A}$ 做了一次相应的初等行(列)变换.

类似地,也可以定义分块初等矩阵,其乘积性质也类同于上.

§4.3 可逆矩阵

定义 4.8 对于矩阵 $\boldsymbol{A}$,若存在矩阵 $\boldsymbol{B}$ 使得 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}$,则称 $\boldsymbol{A}$ 为**可逆矩阵**,而 $\boldsymbol{B}$ 称为 $\boldsymbol{A}$ 的**逆矩阵**.

由定义知,若矩阵 $\boldsymbol{A}$ 可逆,则 $\boldsymbol{A}$ 一定是方阵,同时其逆矩阵是唯一的.事实上,若 $\boldsymbol{B}_1,\boldsymbol{B}_2$ 都是 $\boldsymbol{A}$ 的逆矩阵.则有

$$\boldsymbol{B}_1=\boldsymbol{B}_1\boldsymbol{I}=\boldsymbol{B}_1(\boldsymbol{AB}_2)=(\boldsymbol{B}_1\boldsymbol{A})\boldsymbol{B}_2=\boldsymbol{IB}_2=\boldsymbol{B}_2.$$

下面记可逆矩阵 $\boldsymbol{A}$ 的逆矩阵为 $\boldsymbol{A}^{-1}$.

例 4.6 初等矩阵都可逆,且有 $\boldsymbol{P}(i,j)^{-1}=\boldsymbol{P}(i,j)$,$\boldsymbol{P}(i(c))^{-1}=\boldsymbol{P}(i(c^{-1}))$,$\boldsymbol{P}(j,i(k))^{-1}=\boldsymbol{P}(j,i(-k))$.故初等矩阵的逆仍为初等矩阵.

证明 直接验证即可.

如果 $\boldsymbol{A}$ 可逆,则有 $\boldsymbol{AA}^{-1}=\boldsymbol{I}$,即 $|\boldsymbol{A}|\cdot|\boldsymbol{A}^{-1}|=|\boldsymbol{I}|=1$.所以 $|\boldsymbol{A}|\neq 0$,且 $|\boldsymbol{A}^{-1}|=|\boldsymbol{A}|^{-1}$.即若 $\boldsymbol{A}$ 可逆,则其行列式一定不为 0. 反之如何?为此,引入矩阵的伴随矩阵这个概念. □

定义 4.9 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$,$n\geqslant 2$,定义 $\boldsymbol{A}$ 的**伴随矩阵**为

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}.$$

其中 A_{ij}是 $\boldsymbol{A}$ 中第 i 行第 j 列元素 a_{ij}的代数余子式.

利用行列式按一行及一列展开性质,即定理 3.2,定理 3.2′和命题 3.2.可以得到

$$\boldsymbol{AA}^* = \boldsymbol{A}^*\boldsymbol{A} = |\boldsymbol{A}|\boldsymbol{I},$$

所以若 $|\boldsymbol{A}| \neq 0$,则有

$$\boldsymbol{A}\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^* = \frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I},$$

即 $\boldsymbol{A}$ 可逆且 $\boldsymbol{A}^{-1} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*$.这样我们便证明了如下定理.

定理 4.2 矩阵 $\boldsymbol{A}$ 可逆的充要条件是 $|\boldsymbol{A}| \neq 0$.

注意求 n 阶矩阵的伴随矩阵需要计算 n^2 个 $n-1$ 阶行列式,计算量太大.所以通常不用伴随矩阵的方法求可逆矩阵的逆矩阵.

命题 4.6 若方阵 $\boldsymbol{A}, \boldsymbol{B}$ 满足 $\boldsymbol{AB} = \boldsymbol{I}$(或 $\boldsymbol{BA} = \boldsymbol{I}$),则 $\boldsymbol{A}$ 可逆且 $\boldsymbol{A}^{-1} = \boldsymbol{B}$.

证明 若 $\boldsymbol{AB} = \boldsymbol{I}$.则有 $|\boldsymbol{A}| \cdot |\boldsymbol{B}| = 1$,即 $|\boldsymbol{A}| \neq 0$,所以 $\boldsymbol{A}$ 可逆 .

由于 $\boldsymbol{A}^{-1}$存在,在 $\boldsymbol{AB} = \boldsymbol{I}$ 两边左乘 $\boldsymbol{A}^{-1}$得

$$\boldsymbol{A}^{-1}(\boldsymbol{AB}) = \boldsymbol{A}^{-1}\boldsymbol{I}.$$

即 $\boldsymbol{B} = \boldsymbol{A}^{-1}$. □

命题 4.7 设 $\boldsymbol{A}$ 可逆,$k \neq 0$,则 $\boldsymbol{A}^{-1}, \boldsymbol{A}^{\mathrm{T}}, k\boldsymbol{A}$ 都可逆且有

$$(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}, \quad (\boldsymbol{A}^{\mathrm{T}})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm{T}}, \quad (k\boldsymbol{A})^{-1} = k^{-1}\boldsymbol{A}^{-1}.$$

证明 直接验证即可. □

命题 4.8 设同阶方阵 $\boldsymbol{A}, \boldsymbol{B}$ 都可逆,则 $\boldsymbol{AB}$ 也可逆且$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

证明 由于 $\boldsymbol{A}^{-1}, \boldsymbol{B}^{-1}$都存在,故 $\boldsymbol{B}^{-1}A^{-1}$存在.由

$$(\boldsymbol{AB})(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}) = \boldsymbol{A}(\boldsymbol{BB}^{-1})\boldsymbol{A}^{-1} = \boldsymbol{AIA}^{-1} = \boldsymbol{I}$$

和命题 4.6 知,$\boldsymbol{AB}$ 可逆,且$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$. □

定理 4.3 方阵 $\boldsymbol{A}$ 可逆的充要条件是 $\boldsymbol{A}$ 可以写成初等矩阵的乘积.

证明 充分性.因为初等矩阵都可逆,所以它们的乘积也可逆.

必要性.设 $\boldsymbol{A}$ 可逆,则 $|\boldsymbol{A}| \neq 0$. 故齐次线性方程组 $\boldsymbol{AX} = \boldsymbol{0}$ 只有零解.所以对 $\boldsymbol{A}$ 做初等行变换把 $\boldsymbol{A}$ 化成阶梯形矩阵时没有零行.再继续做初等行变换化成简化的阶梯形矩阵一定是单位矩阵 $\boldsymbol{I}$.由于对矩阵做一次初等行变换相当于在矩阵的左侧乘一个初等矩阵.这表明存在初等矩阵 $\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_t$ 使得

$$\boldsymbol{P}_t \cdots \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{A} = \boldsymbol{I}.$$

所以 $A=(P_t\cdots P_2P_1)^{-1}=P_1^{-1}P_2^{-1}\cdots P_t^{-1}$,为初等矩阵的乘积. □

下面考虑求解矩阵方程 $AX=B$,其中 A 为 n 阶可逆矩阵,B 为 $n\times s$ 矩阵.两边同时左乘 A^{-1} 得 $X=A^{-1}B$,即上面矩阵方程有唯一解. 那么如何有效地把这个唯一解求出来?

由于 A 可逆,A^{-1} 也可逆,A^{-1} 为初等矩阵的乘积. 设 $A^{-1}=Q_k\cdots Q_2Q_1$,构做矩阵 $R=(A\vdots B)$,即把矩阵 B 放在 A 的右端排成的 $n\times(n+s)$ 矩阵,则

$$Q_k\cdots Q_2Q_1R=A^{-1}R=A^{-1}(A\vdots B)=(I\vdots A^{-1}B).$$

这表明当我们对矩阵 R 做初等行变换把它的前 n 列(即 A 这部分)化成单位矩阵时(即把 R 用初等行变换化成简化的阶梯形矩阵时),R 的后 s 列(即 B 这部分)就变成了 $A^{-1}B$. 它就是矩阵方程 $AX=B$ 的唯一解.

例 4.7 设 $A=\begin{pmatrix}1&-2&0\\4&-2&-2\\-3&1&2\end{pmatrix}$,$B=\begin{pmatrix}-1&4\\2&5\\1&-3\end{pmatrix}$,求 $AX=B$ 的解 X.

解

$$R=\begin{pmatrix}1&-2&0&-1&4\\4&-2&-2&2&5\\-3&1&2&1&-3\end{pmatrix}\longrightarrow\begin{pmatrix}1&-2&0&-1&4\\0&6&-2&6&-11\\0&-5&2&-2&9\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&-2&0&-1&4\\0&1&0&4&-2\\0&-5&2&-2&9\end{pmatrix}\longrightarrow\begin{pmatrix}1&-2&0&-1&4\\0&1&0&4&-2\\0&0&2&18&-1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&-2&0&-1&4\\0&1&0&4&-2\\0&0&1&9&-\frac{1}{2}\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&0&7&0\\0&1&0&4&-2\\0&0&1&9&-\frac{1}{2}\end{pmatrix}.$$

所以

$$X=\begin{pmatrix}7&0\\4&-2\\9&-\frac{1}{2}\end{pmatrix}.$$

求可逆矩阵 A 的逆矩阵 A^{-1},即求矩阵方程 $AX=I$ 的解,所以我们可以构造 $n\times2n$ 矩阵 (A,I).对它做初等行变换化成简化的阶梯形矩阵后,它的后 n 列化成的矩阵即为 A^{-1}. 这种方法称为**求可逆矩阵的初等变换法**.

例 4.8 求矩阵

$$A=\begin{pmatrix}2&1&-2\\1&2&2\\2&-2&1\end{pmatrix}$$

的逆矩阵.

解

$$\begin{pmatrix}2&1&-2&1&0&0\\1&2&2&0&1&0\\2&-2&1&0&0&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&2&2&0&1&0\\2&1&-2&1&0&0\\2&-2&1&0&0&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&2&2&0&1&0\\0&-3&-6&1&-2&0\\0&-6&-3&0&-2&1\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&2&2&0&1&0\\0&-3&-6&1&-2&0\\0&0&9&-2&2&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&2&2&0&1&0\\0&1&2&-\frac{1}{3}&\frac{2}{3}&0\\0&0&1&-\frac{2}{9}&\frac{2}{9}&\frac{1}{9}\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1&2&0&\frac{4}{9}&\frac{5}{9}&-\frac{2}{9}\\0&1&0&\frac{1}{9}&\frac{2}{9}&-\frac{2}{9}\\0&0&1&-\frac{2}{9}&\frac{2}{9}&\frac{1}{9}\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&0&\frac{2}{9}&\frac{1}{9}&\frac{2}{9}\\0&1&0&\frac{1}{9}&\frac{2}{9}&-\frac{2}{9}\\0&0&1&-\frac{2}{9}&\frac{2}{9}&\frac{1}{9}\end{pmatrix},$$

所以

$$\boldsymbol{A}^{-1}=\begin{pmatrix}\frac{2}{9}&\frac{1}{9}&\frac{2}{9}\\\frac{1}{9}&\frac{2}{9}&-\frac{2}{9}\\-\frac{2}{9}&\frac{2}{9}&\frac{1}{9}\end{pmatrix}.$$

显然,当 n 较大时,用初等变换法求逆矩阵比用伴随矩阵求逆要有效得多.

形如

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11}&\boldsymbol{O}&\cdots&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{A}_{22}&\cdots&\boldsymbol{O}\\\vdots&\vdots&&\vdots\\\boldsymbol{O}&\boldsymbol{O}&\cdots&\boldsymbol{A}_{ss}\end{pmatrix}$$

的分块矩阵称为**分块对角矩阵**,其中 $\boldsymbol{A}_{ii}$ 为 n_i 阶方阵,$1\leqslant i\leqslant s$,显然 $|\boldsymbol{A}|=|\boldsymbol{A}_{11}|\cdot|\boldsymbol{A}_{22}|\cdot\cdots\cdot|\boldsymbol{A}_{ss}|$,所以矩阵 $\boldsymbol{A}$ 可逆当且仅当每个小块 $\boldsymbol{A}_{ii}$ 都可逆,并且容易验证

$$A^{-1}=\begin{pmatrix}A_{11}^{-1} & O & \cdots & O\\ O & A_{22}^{-1} & \cdots & O\\ \vdots & \vdots & & \vdots\\ O & O & \cdots & A_{ss}^{-1}\end{pmatrix}$$

仍为分块对角矩阵．类似地可以定义**分块上三角矩阵**

$$\begin{pmatrix}A_{11} & A_{12} & \cdots & A_{1s}\\ O & A_{22} & \cdots & A_{2s}\\ \vdots & \vdots & & \vdots\\ O & O & \cdots & A_{ss}\end{pmatrix}$$

和**分块下三角矩阵**

$$\begin{pmatrix}A_{11} & O & \cdots & O\\ A_{21} & A_{22} & \cdots & O\\ \vdots & \vdots & & \vdots\\ A_{s1} & A_{s2} & \cdots & A_{ss}\end{pmatrix},$$

其中主对角线上的块 A_{ii} 为 n_i 阶方阵．同样易知这样的分块矩阵可逆当且仅当其主对角线上的块 A_{ii} 都可逆，$1\leqslant i\leqslant s$. 这时，如何求它的逆呢？下面看最简单的情形．

$$A=\begin{pmatrix}A_{11} & A_{12}\\ O & A_{22}\end{pmatrix},$$

其中 A_{11}, A_{22} 分别为 n_1 和 n_2 阶可逆矩阵，A_{12} 为 $n_1\times n_2$ 矩阵．由于

$$\begin{pmatrix}I_{n_1} & -A_{12}A_{22}^{-1}\\ O & I_{n_2}\end{pmatrix}\begin{pmatrix}A_{11} & A_{12}\\ O & A_{22}\end{pmatrix}=\begin{pmatrix}A_{11} & O\\ O & A_{22}\end{pmatrix},$$

所以

$$A=\begin{pmatrix}A_{11} & A_{12}\\ O & A_{22}\end{pmatrix}=\begin{pmatrix}I_{n_1} & -A_{12}A_{22}^{-1}\\ O & I_{n_2}\end{pmatrix}^{-1}\begin{pmatrix}A_{11} & O\\ O & A_{22}\end{pmatrix},$$

故

$$A^{-1}=\begin{pmatrix}A_{11} & O\\ O & A_{22}\end{pmatrix}^{-1}\begin{pmatrix}I_{n_1} & -A_{12}A_{22}^{-1}\\ O & I_{n_2}\end{pmatrix}$$

$$=\begin{pmatrix}A_{11}^{-1} & O\\ O & A_{22}^{-1}\end{pmatrix}\begin{pmatrix}I_{n_1} & -A_{12}A_{22}^{-1}\\ O & I_{n_2}\end{pmatrix}=\begin{pmatrix}A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1}\\ O & A_{22}^{-1}\end{pmatrix},$$

它仍为分块上三角矩阵，且主对角线上的块为原矩阵对应的主对角线上块的逆．

设有 n 个未知量、n 个方程的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \tag{4.2}$$

的系数矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 的行列式不为 0. 此方程组可写成矩阵形式 $\boldsymbol{AX}=\boldsymbol{\beta}$，其中 $\boldsymbol{X}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$，$\boldsymbol{\beta}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$. 由于 $|\boldsymbol{A}|\neq 0$，所以 $\boldsymbol{A}$ 可逆，故此方程组有唯一解 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{\beta}$. 由 $\boldsymbol{A}^{-1}$ 的伴随矩阵形式

$$\boldsymbol{X}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*\boldsymbol{\beta}=\frac{1}{|\boldsymbol{A}|}\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$=\frac{1}{|\boldsymbol{A}|}\begin{pmatrix} b_1A_{11}+b_2A_{21}+\cdots+b_nA_{n1} \\ b_1A_{12}+b_2A_{22}+\cdots+b_nA_{n2} \\ \vdots \\ b_1A_{1n}+b_2A_{2n}+\cdots+b_nA_{nn} \end{pmatrix}=\frac{1}{|\boldsymbol{A}|}\begin{pmatrix} |\boldsymbol{B}_1| \\ |\boldsymbol{B}_2| \\ \vdots \\ |\boldsymbol{B}_n| \end{pmatrix},$$

其中 $\boldsymbol{B}_j$ 是把 $\boldsymbol{A}$ 的第 j 列换成 $\boldsymbol{\beta}$ 这一列所得到的矩阵. 这表明线性方程组(4.2)有唯一解

$$x_1=\frac{|\boldsymbol{B}_1|}{|\boldsymbol{A}|},\quad x_2=\frac{|\boldsymbol{B}_2|}{|\boldsymbol{A}|},\quad \cdots,\quad x_n=\frac{|\boldsymbol{B}_n|}{|\boldsymbol{A}|},$$

这便是克拉默法则.

习 题 四

1. 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 和 $\boldsymbol{I}$ 都是 n 阶方阵，下列等式或结论是否成立？

(1) $(\boldsymbol{A}+\boldsymbol{B})^2=\boldsymbol{A}^2+2\boldsymbol{AB}+\boldsymbol{B}^2$；

(2) 若 $\boldsymbol{A}^2=\boldsymbol{A}$，则 $\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{A}=\boldsymbol{I}$；

(3) 若 $\boldsymbol{A}$ 为对称矩阵，则 $\boldsymbol{A}^k$ 也为对称矩阵(k 为正整数).

2. 设 $\boldsymbol{A}=\begin{pmatrix}1&3&0\\2&-2&1\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}1&-1&-2\\0&2&-1\end{pmatrix}$, 且 $4(\boldsymbol{X}-\boldsymbol{A})+\boldsymbol{B}=2(\boldsymbol{X}+\boldsymbol{B})-3\boldsymbol{A}$, 求 $\boldsymbol{X}$.

3. 设 $\boldsymbol{A}=\begin{pmatrix}1&2\\x&-1\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}1&y\\1&0\end{pmatrix}$ 且 $\boldsymbol{AB}=\boldsymbol{BA}$, 求 x 和 y.

4. 计算下列矩阵的乘积:

(1) $\begin{pmatrix}1&2&-1\\-2&1&0\\1&0&3\end{pmatrix}\begin{pmatrix}3&3\\1&-1\\2&4\end{pmatrix}$;　　(2) $\begin{pmatrix}1&3&0\\-2&-1&1\end{pmatrix}\begin{pmatrix}1&3&-1&0\\0&-1&2&1\\2&4&0&1\end{pmatrix}$;

(3) $(x_1,x_2,x_3)\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{12}&a_{22}&a_{23}\\a_{13}&a_{23}&a_{33}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$;　　(4) $\begin{pmatrix}1&1\\0&1\end{pmatrix}^n$, 其中 n 为正整数.

5. 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 n 阶方阵.求证:

$$(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})=\boldsymbol{A}^2-\boldsymbol{B}^2 \Leftrightarrow \boldsymbol{AB}=\boldsymbol{BA}.$$

6. 求与 $\boldsymbol{A}=\begin{pmatrix}1&1\\0&1\end{pmatrix}$ 可交换的所有矩阵.

7. 在下列矩阵方程中求矩阵 $\boldsymbol{X}$:

(1) $\boldsymbol{X}\begin{pmatrix}1&2\\3&4\end{pmatrix}=\begin{pmatrix}3&5\\5&9\end{pmatrix}$;

(2) $\begin{pmatrix}1&2&-3\\2&2&-4\\2&-1&0\end{pmatrix}\boldsymbol{X}=\begin{pmatrix}1&-3&0\\10&2&7\\10&7&8\end{pmatrix}$.

8. 设方阵 $\boldsymbol{A}$ 满足矩阵方程 $\boldsymbol{A}^2-2\boldsymbol{A}+5\boldsymbol{I}=\boldsymbol{O}$, 求证: $\boldsymbol{A}$ 可逆, 并求 $\boldsymbol{A}^{-1}$.

9. 求下列矩阵的逆矩阵:

(1) $\boldsymbol{A}=\begin{pmatrix}2&-1&0&0\\-3&2&0&0\\0&0&3&-4\\0&0&-2&3\end{pmatrix}$;　　(2) $\boldsymbol{A}=\begin{pmatrix}0&0&1&2\\0&0&2&1\\2&1&0&0\\1&3&0&0\end{pmatrix}$.

10. 已知 $\boldsymbol{A},\boldsymbol{B}$ 为 3 阶矩阵, 且满足 $2\boldsymbol{A}^{-1}\boldsymbol{B}=\boldsymbol{B}-4\boldsymbol{I}$, 其中 $\boldsymbol{I}$ 是 3 阶单位矩阵.

(1) 求证: 矩阵 $\boldsymbol{A}-2\boldsymbol{I}$ 可逆;

(2) 若 $\boldsymbol{B}=\begin{pmatrix}1&-2&0\\1&2&0\\0&0&2\end{pmatrix}$, 求矩阵 $\boldsymbol{A}$.

11. 设 $\boldsymbol{A}=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&-1\end{pmatrix}$, $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}$, 其中 $\boldsymbol{P}$ 为 3 阶可逆矩阵, 求 $\boldsymbol{B}^{2\,016}-2\boldsymbol{A}^2$.

第五章 向量空间

§5.1 n 维向量空间

在第一章中,我们在三维空间$\mathbb{R}^3$里,介绍了空间向量及其线性运算.进一步用之分析了直线、平面、角度、距离等一系列几何问题与物理现象.然而,向量的概念并不仅仅存在于三维的几何空间中.例如研究人造卫星在太空运行时的状态,人们感兴趣的不只是它的轨迹,还希望知道在某个时刻它所处的位置、其表面温度、压力等物理参数.这时,我们需要用六元数组(t,x,y,z,e,p)来表示.因此,有必要拓展向量空间的概念.引入由 n 元数组构成的 n 维向量,并抽象出向量空间的概念.

本章不仅要讨论一般向量组的概念,还要建立矩阵的秩的概念,最后以此为背景完整地解决一般线性方程组的解的相关理论.

本章中用 F 表示数域,例如:$F=\mathbb{R}$(实数域)和 $F=\mathbb{C}$(复数域).

定义 5.1 n 个数 $a_1,a_2,\cdots,a_n\in F$ 组成的有序数组 $\boldsymbol{\alpha}$ 称为数域 F 上的 n **维向量**(vector),记作

$$\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)\text{ 或 }\boldsymbol{\alpha}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},$$

其中 a_i 称为向量 $\boldsymbol{\alpha}$ 的第 $i(i=1,2,\cdots,n)$ **分量**(component).前一个表示式称为**行向量**,后一个称为**列向量**.

用矩阵转置的记号,列向量也记作 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$.本书以后一般都采用列向量表示.另外,把零向量记为 $\mathbf{0}=(0,0,\cdots,0)^{\mathrm{T}}$.

下面来规定向量相等以及向量的线性运算(加法、数乘).

定义 5.2 两个 n 维向量 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$ 与 $\boldsymbol{\beta}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$ 称为相等的向量当且仅当对应分量全相等，即 $a_i=b_i(i=1,2,\cdots,n)$.

定义向量的加法与数乘如下：

定义 5.3 $\boldsymbol{\alpha}+\boldsymbol{\beta}=(a_1+b_1,a_2+b_2,\cdots,a_n+b_n)^{\mathrm{T}}$,

$k\boldsymbol{\alpha}=(ka_1,ka_2,\cdots,ka_n)^{\mathrm{T}}$.

实际上，n 维行向量可视为 $1\times n$ 矩阵，n 维列向量可视为 $n\times 1$ 矩阵，故向量的线性运算可由矩阵的线性运算直接导出.

在本节最后，给出 n 维向量的几个简单性质（设 $\boldsymbol{\alpha}$ 为 n 维向量，k 为 F 中的一个数）：

(1) $0\boldsymbol{\alpha}=\mathbf{0}$;

(2) $k\mathbf{0}=\mathbf{0}$;

(3) 若 $k\boldsymbol{\alpha}=\mathbf{0}$，则 $k=0$ 或 $\boldsymbol{\alpha}=\mathbf{0}$.

5.1.1 n 维向量空间与子空间

定义 5.4 数域 F 上的全体 n 维向量的集合，对向量的加法和数乘两种运算，构成数域 F 上的 **n 维向量空间**，记作 F^n.

当 $n=2,3$ 且 $F=\mathbb{R}$ 时，就是我们熟悉的实平面与几何空间.

我们知道，集合有子集合的概念，类似地，可给出子空间的概念.

定义 5.5 设 W 是 F^n 的一个非空子集合，满足：

(1) 对向量的加法是封闭的：$\forall\, \boldsymbol{\alpha},\boldsymbol{\beta}\in W$ 有 $\boldsymbol{\alpha}+\boldsymbol{\beta}\in W$;

(2) 对向量的数乘是封闭的：$\forall\, \boldsymbol{\alpha}\in W$ 和 $\forall\, k\in F$ 有 $k\boldsymbol{\alpha}\in W$,

则称 W 是 F^n 的一个**子空间**.

例如，设 $V_1=\{(x,0,\cdots,0)^{\mathrm{T}}\mid x\in F\}$. 显然 $V_1\subset F^n$，对向量的加法和数乘运算都是封闭的，所以 V_1 是 F^n 的一个子空间.

又如，只有一个零向量的集合：$V_0=\{(0,0,\cdots,0)^{\mathrm{T}}\}$，对向量的加法和数乘运算都是封闭的，所以 V_0 是 F^n 的一个子空间. 此子空间称为**零子空间**.

5.1.2 向量组的线性组合与线性表出

在几何空间 $\mathbb{R}^3$ 中，在自然基 $\boldsymbol{i}=(1,0,0)^{\mathrm{T}},\boldsymbol{j}=(0,1,0)^{\mathrm{T}},\boldsymbol{k}=(0,0,1)^{\mathrm{T}}$ 下，任意一个向量 $\boldsymbol{\alpha}=(x,y,z)^{\mathrm{T}}$ 可表示为 $\boldsymbol{\alpha}=x\boldsymbol{i}+y\boldsymbol{j}+z\boldsymbol{k}$.

下面，我们将这样的概念推广到 n 维向量空间 F^n 中来.

定义 5.6 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是 n 维向量，$k_1,k_2,\cdots,k_s\in F$，则称

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s$$

是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的一个（F-）**线性组合**.

如果对向量$\boldsymbol{\beta}$,存在一组实数$k_1,k_2,\cdots,k_s$,使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s,$$

则称$\boldsymbol{\beta}$是$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$的一个线性组合,也称$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$线性表出.

例如,对$\boldsymbol{\alpha}_1=(1,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,1)^{\mathrm{T}},\boldsymbol{\beta}=(-3,2)^{\mathrm{T}}$,有$\boldsymbol{\beta}=-5\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2$.即$\boldsymbol{\beta}$是$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$的一个线性组合.

又如,对$\boldsymbol{\alpha}_1=(0,1)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,2)^{\mathrm{T}},\boldsymbol{\alpha}_3=(-2,4)^{\mathrm{T}},\boldsymbol{\beta}=(3,5)^{\mathrm{T}}$,有$\boldsymbol{\beta}=-\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2+0\boldsymbol{\alpha}_3$,以及$\boldsymbol{\beta}=7\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3$,还有$\boldsymbol{\beta}=11\boldsymbol{\alpha}_1+0\boldsymbol{\alpha}_2-\dfrac{3}{2}\boldsymbol{\alpha}_3$等等,故$\boldsymbol{\beta}$是$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$的线性组合,而且线性表出的方式不是唯一的.

再如,对$\boldsymbol{\alpha}_1=(2,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(3,0)^{\mathrm{T}},\boldsymbol{\beta}=(0,1)^{\mathrm{T}}$,无论$k_1,k_2$取什么数,$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\neq\boldsymbol{\beta}$,即$\boldsymbol{\beta}$不能由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$线性表出.

例 5.1 已知$\boldsymbol{\alpha}_1=(-1,0,1,2)^{\mathrm{T}},\boldsymbol{\alpha}_2=(3,4,-2,5)^{\mathrm{T}},\boldsymbol{\alpha}_3=(1,4,0,9)^{\mathrm{T}},\boldsymbol{\beta}=(5,4,-4,1)^{\mathrm{T}}$.试问$\boldsymbol{\beta}$能否由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性表出?如果能线性表出就写出其表达式.

解 设$\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3$,按分量写出来即有

$$\begin{cases}-x_1+3x_2+x_3=5,\\ 4x_2+4x_3=4,\\ x_1-2x_2=-4,\\ 2x_1+5x_2+9x_3=1.\end{cases}$$

由此知$\boldsymbol{\beta}$能否由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性表出,等价于问上述线性方程组是否有解.

用高斯消元法得

$$\left(\begin{array}{ccc|c}-1&3&1&5\\0&4&4&4\\1&-2&0&-4\\2&5&9&1\end{array}\right)\to\left(\begin{array}{ccc|c}-1&3&1&5\\0&1&1&1\\0&1&1&1\\0&11&11&11\end{array}\right)\to\left(\begin{array}{ccc|c}-1&3&1&5\\0&1&1&1\\0&0&0&0\\0&0&0&0\end{array}\right)\to\left(\begin{array}{ccc|c}1&0&2&-2\\0&1&1&1\\0&0&0&0\\0&0&0&0\end{array}\right),$$

此线性方程组有解:$x_1=-2-2t,x_2=1-t,x_3=t$.

也就是$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$线性表出,因为$t$可取任意值,因此表示方法有无穷多种,表达式为

$$\boldsymbol{\beta}=(-2-2t)\boldsymbol{\alpha}_1+(1-t)\boldsymbol{\alpha}_2+t\boldsymbol{\alpha}_3.$$

由线性组合可构造子空间.例如,$\boldsymbol{\alpha}_1=(-1,0,1,2)^{\mathrm{T}},\boldsymbol{\alpha}_2=(3,4,-2,5)^{\mathrm{T}},\boldsymbol{\alpha}_3=(1,4,0,9)^{\mathrm{T}}$,令

$$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=\{k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\alpha}_3\mid k_1,k_2,k_3\in F\},$$

易见$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$对向量加法和数乘是封闭的.由子空间定义,$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$是$F^4$的一个子空间,称为由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$生成的子空间.类似地,可定义由任一向量组生成的子空间.

如果给定一个 $m\times n$ 矩阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$，其中每个 $\boldsymbol{\alpha}_i(i=1,2,\cdots,n)$ 是一个 m 维列向量.令

$$L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=\{k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n \mid k_1,k_2,\cdots,k_n\in F\},$$

则 $L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 是由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 生成的子空间，称为矩阵 $\boldsymbol{A}$ 的**列空间**，它是 F^m 的子空间.类似地，可由 $\boldsymbol{A}$ 的 m 个行向量生成 F^n 的一个子空间，称为 $\boldsymbol{A}$ 的**行空间**.

§ 5.2 向量组的线性相关性

在第一章中，我们讨论过几何空间 $\mathbb{R}^3$ 中，向量的共线与共面问题.具体地：若两个向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 共线 $\Leftrightarrow$ 存在实数 $\lambda\in\mathbb{R}$，使得 $\boldsymbol{\alpha}=\lambda\boldsymbol{\beta}$ 或 $\boldsymbol{\beta}=\lambda\boldsymbol{\alpha}$.

经过一定变形后：若两个向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 共线 $\Leftrightarrow$ 存在不全为 0 的实数 k_1,k_2，使得 $k_1\boldsymbol{\alpha}+k_2\boldsymbol{\beta}=\mathbf{0}$.

若三个向量共面 $\Leftrightarrow\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 中某一个可被另外两个线性表出 $\Leftrightarrow$ 存在不全为 0 的实数 $\lambda_1,\lambda_2,\lambda_3$，使得 $\lambda_1\boldsymbol{\alpha}+\lambda_2\boldsymbol{\beta}+\lambda_3\boldsymbol{\gamma}=\mathbf{0}$.

下面，我们将共线，共面的概念推广到一般 n 维空间中.此时，几何意义很难想象，所以主要从代数的角度来刻画.

定义 5.7 给定向量空间 F^n 中 s 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$，如果存在不全为 0 的 $k_1,k_2,\cdots,k_s\in F$，使得 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s=\mathbf{0}$，则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$（$F-$）**线性相关**（linear dependence）.否则称 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$（$F-$）**线性无关**（linear independence）.

由定义 5.7，在 $\mathbb{R}^3$ 中，共线的向量及共面的向量都是线性相关的.

例 $\boldsymbol{\alpha}_1=\begin{pmatrix}1\\2\\1\end{pmatrix},\boldsymbol{\alpha}_2=\begin{pmatrix}2\\4\\2\end{pmatrix},\boldsymbol{\alpha}_3=\begin{pmatrix}1\\3\\5\end{pmatrix}$ 是 $\mathbb{R}^3$ 中三个向量，由于 $\boldsymbol{\alpha}_2=2\boldsymbol{\alpha}_1$，因而有 $2\boldsymbol{\alpha}_1+(-1)\boldsymbol{\alpha}_2+0\boldsymbol{\alpha}_3=\mathbf{0}$，系数 $2,-1,0$ 不全为 0，按定义 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关.

由定义，很容易得到以下几条判则：

（1）$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关 $\Leftrightarrow x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_s\boldsymbol{\alpha}_s=\mathbf{0}$ 只有零解.

（2）若向量组含有零向量，则该向量组线性相关.

（3）对于只有一个向量的向量组，有 $\boldsymbol{\alpha}$ 线性无关 $\Leftrightarrow\boldsymbol{\alpha}\neq\mathbf{0}$.

（4）$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s(s\geqslant 2)$ 线性相关 $\Leftrightarrow$ 其中一个向量可由其他向量线性表出.

证明 $\Rightarrow$ 已知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关.由定义 5.7 知存在不全为 0 的数 $k_1,k_2,\cdots,k_s\in F$ 使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s=\mathbf{0},$$

不失一般性,可设 $k_1 \neq 0$,则得

$$\boldsymbol{\alpha}_1 = -\frac{k_2}{k_1}\boldsymbol{\alpha}_2 - \cdots - \frac{k_s}{k_1}\boldsymbol{\alpha}_s,$$

所以 $\boldsymbol{\alpha}_1$ 可由其他向量线性表出.

$\Leftarrow$设其中一个向量可由其他向量线性表出,不失一般性,设 $\boldsymbol{\alpha}_1$ 可由其他向量线性表出,$\boldsymbol{\alpha}_1 = q_2\boldsymbol{\alpha}_2 + \cdots + q_s\boldsymbol{\alpha}_s$,于是有

$$\boldsymbol{\alpha}_1 - q_2\boldsymbol{\alpha}_2 - \cdots - q_s\boldsymbol{\alpha}_s = \mathbf{0},$$

至少 $\boldsymbol{\alpha}_1$ 的系数不为 0,由定义 5.7,$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 线性相关. □

如果给定几个向量,如何判断它们是否线性相关呢?我们来看以下例子.

例 5.2 设 $\boldsymbol{\alpha}_1 = (1,0,0,0)^{\mathrm{T}}, \boldsymbol{\alpha}_2 = (1,1,0,0)^{\mathrm{T}}, \boldsymbol{\alpha}_3 = (1,1,1,0)^{\mathrm{T}}$. 判断它们是否线性相关.

解 设 $k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + k_3\boldsymbol{\alpha}_3 = \mathbf{0}$,得线性方程组

$$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)\begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \mathbf{0},$$

由于$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$已是阶梯阵,所以只有零解:$k_1 = k_2 = k_3 = 0$,因而 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性无关.

因而判断向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 是否线性相关的基本方法是解线性方程组

$$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s)\boldsymbol{X} = \mathbf{0}$$

是否有非零解,于是我们得到以下结果.

(5) $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 线性相关$\Leftrightarrow$方程组$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s)\boldsymbol{X} = \mathbf{0}$ 有非零解.

注 实际上,性质(5)是性质(1)的等价的逆否命题,它们给出了用齐次线性方程组来判定线性相关性的方法.

根据性质(1)与(5),特别地有以下两条:

(6) $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性相关$\Leftrightarrow \det(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) = 0$.

($\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性无关$\Leftrightarrow \det(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) \neq 0$.)

(7) “阶梯形”向量组

$$\boldsymbol{\alpha}_1=\begin{pmatrix}a_{11}\\0\\0\\\vdots\\0\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}a_{12}\\a_{22}\\0\\\vdots\\0\end{pmatrix},\cdots,\boldsymbol{\alpha}_s=\begin{pmatrix}a_{1s}\\a_{2s}\\\vdots\\a_{ss}\\\vdots\\0\end{pmatrix}\quad (a_{ii}\neq 0,i=1,2,\cdots,s)$$

线性无关.

注 性质(7)是例 5.2 的一般化结论.

下面一些结果,有些比较简单,请读者自己证明.

(8) 如果一个向量组中部分向量线性相关,则整个向量组也线性相关.反之,如果一个向量组线性无关,则其任意部分向量组也线性无关.

(9) 如果一个向量组的向量个数大于向量维数,则该向量组必线性相关.或说在 F^n 中一组向量的个数大于 n,则该向量组必线性相关.

(10) 如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关,但 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}$ 线性相关,则 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出,而且表达式是唯一的.

证明 由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}$ 线性相关,因而有不全为 0 的数 $k_1,k_2,\cdots,k_s,k_{s+1}$ 使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s+k_{s+1}\boldsymbol{\beta}=\mathbf{0},$$

可证必有 $k_{s+1}\neq 0$:反证法.如果 $k_{s+1}=0$,则有不全为 0 的实数 $k_1,k_2,\cdots,k_s$ 使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s=\mathbf{0},$$

因而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关,与已知条件矛盾.所以 $k_{s+1}\neq 0$,于是有

$$\boldsymbol{\beta}=-\left(\frac{k_1}{k_{s+1}}\boldsymbol{\alpha}_1+\frac{k_2}{k_{s+1}}\boldsymbol{\alpha}_2+\cdots+\frac{k_s}{k_{s+1}}\boldsymbol{\alpha}_s\right),$$

所以 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出.

唯一性留作思考题. □

(11) 如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关,把每个向量增加相同数目相同位置的分量,则得到的向量组仍然线性无关.反之,如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关,把每个向量去掉相同位置的分量,则得到的向量组仍然线性相关.

证明 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关,他们的分量为

$$\boldsymbol{\alpha}_j=(a_{1j},a_{2j},\cdots,a_{mj})^{\mathrm{T}},\quad j=1,2,\cdots,s.$$

现把每一个向量任意增加 t 个分量,不妨设都增加在最后 t 个位置上,变成

$$\boldsymbol{\beta}_j=(a_{1j},a_{2j},\cdots,a_{mj},a_{(m+1)j},\cdots,a_{(m+t)j})^{\mathrm{T}},\quad j=1,2,\cdots,s.$$

考虑方程组

$$(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s)X=\mathbf{0}, \tag{5.1}$$

由于方程组

$$(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s)X=\mathbf{0} \tag{5.2}$$

是由方程组(5.1)的前 m 个方程组成,故(5.1)的解必是(5.2)的解.如果(5.1)有非零解,则(5.2)也有非零解,这与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关矛盾,所以齐次线性方程组(5.1)只有零解,因而 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性无关.

再证明第 2 个结论:设 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性相关.把每个分量的后 t 个分量去掉,则得到向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$.现考虑方程组(5.1),它有非零解 $\boldsymbol{X}_0\neq\mathbf{0}$.因为(5.1)的解必是(5.2)的解,所以 $\boldsymbol{X}_0$ 也是(5.2)的非零解,因而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关. □

这个性质的理解可从线性方程组的解入手,增加分量相当于增加方程数,于是原来只有零解,更只有零解了.而减少分量相当于减少方程数,于是原来有非零解,更有非零解了.

以上的判则非常重要,作为有关向量组线性相关的常识,必须记住.记住的方法不是通过死记硬背,而是对每一个结论进行证明,同时,通过证明我们对线性相关的概念逐渐熟悉.

例 5.3 判断 $\boldsymbol{\alpha}_1=(1,-1,2,3)^{\mathrm{T}},\boldsymbol{\alpha}_2=(2,2,0,-1)^{\mathrm{T}},\boldsymbol{\alpha}_3=(0,2,5,8)^{\mathrm{T}},\boldsymbol{\alpha}_4=(-1,7,-1,-2)^{\mathrm{T}}$ 的线性相关性.

解 由性质(6),考虑系数行列式

$$\begin{vmatrix}1&2&0&-1\\-1&2&2&7\\2&0&5&-1\\3&-1&8&-2\end{vmatrix}=\begin{vmatrix}1&0&0&0\\-1&4&2&6\\2&-4&5&1\\3&-7&8&1\end{vmatrix}=\begin{vmatrix}4&2&6\\-4&5&1\\-7&8&1\end{vmatrix}=0.$$

所以 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性相关.

例 5.4 已知 $\boldsymbol{\alpha}_1=(1,2,-1,4)^{\mathrm{T}},\boldsymbol{\alpha}_2=(0,-1,a,3)^{\mathrm{T}},\boldsymbol{\alpha}_3=(2,5,3,5)^{\mathrm{T}}$ 线性无关,求 a.

解 考察齐次线性方程组 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3=\mathbf{0}$,用高斯消元法解此齐次线性方程组,得

$$\begin{pmatrix}1&0&2\\2&-1&5\\-1&a&3\\4&3&5\end{pmatrix}\to\begin{pmatrix}1&0&2\\0&-1&1\\0&a&5\\0&3&-3\end{pmatrix}\to\begin{pmatrix}1&0&2\\0&-1&1\\0&0&5+a\\0&0&0\end{pmatrix}.$$

$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关,要求齐次线性方程组只有零解,故 $a\neq-5$.

§5.3 向量组的秩

5.3.1 两个向量组的关系

在§5.1节中,我们已经给出了一个向量被一个向量组线性表出的概念.下面,我们将其推广到一个向量组被另一个向量组线性表出的情形.

定义 5.8 设有两个 n 维向量组

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s, \quad ①$$

$$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t, \quad ②$$

如果向量组①中每个向量 $\boldsymbol{\alpha}_i(i=1,2,\cdots,s)$ 都可由向量组②中的向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性表出,则称**向量组①可由向量组②线性表出**.

当一个向量组可被另一个向量组线性表出时,它们的线性相关性有何关系呢?

定理 5.1 设向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出,且 $t>s$,则 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性相关.

注 (1) 直观上想,以少生多,必线性相关.

(2) 等价的逆否命题:若 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性无关,则 $t\leqslant s$.直观上想:线性无关向量组只能由数量更多的向量组生成.

证明 首先把向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的关系表示出来.由于 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出,于是可设

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)\boldsymbol{A}_{s\times t}.$$

考虑 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 的线性相关性,等价于考虑方程组

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)\boldsymbol{X}=\mathbf{0}.$$

综合上两式有

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)\boldsymbol{A}_{s\times t}\boldsymbol{X}=\mathbf{0}.$$

由于矩阵 $\boldsymbol{A}_{s\times t}$ 的行数小于列数,方程组 $\boldsymbol{A}_{s\times t}\boldsymbol{X}=\mathbf{0}$ 必有非零解,且该非零解也是上述方程组的解.从而知 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性相关. □

如果向量组①、②可以互相线性表出,则称这**两个向量组等价**(equivalence),记为①~②.

等价的两个向量组的线性相关性也存在某种联系,究竟是什么样的联系呢? 为此,我们先要分析表达一个向量组的本质的东西是什么.

5.3.2　极大线性无关组

定义 5.9　在向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 中，若存在 r 个向量 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性无关，再加进任意一个向量 $\boldsymbol{\alpha}_j(j=1,2,\cdots,s)$ 就线性相关，则称 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 是向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的一个**极大线性无关组**.

例如，向量组 $\boldsymbol{\alpha}_1=\begin{pmatrix}1\\0\end{pmatrix},\boldsymbol{\alpha}_2=\begin{pmatrix}1\\2\end{pmatrix},\boldsymbol{\alpha}_3=\begin{pmatrix}2\\3\end{pmatrix}$ 中，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是线性无关的，而 $\boldsymbol{\alpha}_3=\dfrac{1}{2}\boldsymbol{\alpha}_1+\dfrac{3}{2}\boldsymbol{\alpha}_2$，即添加 $\boldsymbol{\alpha}_3$ 后就线性相关，所以，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 就是一个极大线性无关组.

实际上，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3$ 与 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 也都是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 的极大线性无关组.这说明，一个向量组的极大线性无关组往往是不唯一的，我们将在下一小节讨论它们之间有什么关系.

下面介绍求极大线性无关组的方法.其原理是用解方程组的方法.

以向量组为列向量构成系数矩阵，得到以下的齐次线性方程组：

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)\boldsymbol{X}=\boldsymbol{0}.$$

对上式的系数矩阵进行高斯消元法：

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)\rightarrow\begin{pmatrix}c_{11} & \cdots & c_{1r} & \cdots & c_{1s}\\ 0 & \ddots & \vdots & & \vdots\\ \vdots & \cdots & c_{rr} & \cdots & c_{rs}\\ 0 & \cdots & 0 & \cdots & 0\\ \vdots & & \vdots & & \vdots\\ 0 & \cdots & 0 & \cdots & 0\end{pmatrix},$$

其中 $c_{ii}\neq 0,i=1,\cdots,r$.由此可见，$c_{ii}$ 所对应的向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是线性无关的，加入任意其他的向量，方程组有非零解，因而线性相关.所以 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是一组极大无关组.注意：用高斯消元法得到的是一般的阶梯形（主元素不一定在前 r 列）.因此极大无关组可由阶梯形来得到.

例 5.5　给定向量组：$\boldsymbol{\alpha}_1=(1,-1,2,3)^{\mathrm{T}},\boldsymbol{\alpha}_2=(0,2,5,8)^{\mathrm{T}},\boldsymbol{\alpha}_3=(2,2,0,-1)^{\mathrm{T}},\boldsymbol{\alpha}_4=(-1,7,-1,-2)^{\mathrm{T}}$.求它的极大线性无关组，并把其余向量用极大无关组线性表出.

解　考虑方程组 $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)\boldsymbol{X}=\boldsymbol{0}$.

对系数矩阵进行高斯消元法，可得：

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)\rightarrow\begin{pmatrix}1 & 0 & 2 & -1\\ -1 & 2 & 2 & 7\\ 2 & 5 & 0 & -1\\ 3 & 8 & -1 & -2\end{pmatrix}\xrightarrow{\text{（省略中间过程）}}\cdots\rightarrow\begin{pmatrix}1 & 0 & 2 & -1\\ 0 & 1 & 2 & 3\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0\end{pmatrix}$$

所以 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的一个极大无关组为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$.

把其余向量用极大无关组线性表出,就是取自由变量 $x_4=-1$,

解得 $x_3=1,x_2=1,x_1=-3$.所以得

$$\boldsymbol{\alpha}_4=-3\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3.$$

小结:列向量消去法的解题步骤:

(1) 把向量组用列向量表成矩阵:$\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)$.

(2) 对 $\boldsymbol{A}$ 用高斯消元法消为阶梯形.

(3) 阶梯形的非 0 行主元素所在的列对应的向量组就是极大线性无关组.

(4) 如果要把其余向量用极大线性无关组表出,方法如下:

分别取某个自由变量为 1,其余自由变量为 0,得到齐次方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的一组非零解,从而得到对应向量用极大无关组的一个线性表出.由上节性质(10)知,对固定的极大无关组,该线性表出是唯一的.

5.3.3 向量组的秩

由上一小节的例子,及高斯消元法可有不同做法,我们知道:一个向量组的极大线性无关组的选取方法可以不是唯一的,但我们将要说明不同极大无关组的个数是相同的,也就是说极大无关组的个数是一个反映给定向量组的线性相关性的本质的不变量.

定理 5.2 若 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 与 $\boldsymbol{\alpha}_{j_1},\boldsymbol{\alpha}_{j_2},\cdots,\boldsymbol{\alpha}_{j_t}$ 都是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的极大线性无关组,则 $r=t$.

证明 由定义 5.9 知,$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 与 $\boldsymbol{\alpha}_{j_1},\boldsymbol{\alpha}_{j_2},\cdots,\boldsymbol{\alpha}_{j_t}$ 可相互线性表出,而 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性无关,由定理 5.1 的等价命题知 $r\leqslant t$.同理,由 $\boldsymbol{\alpha}_{j_1},\cdots,\boldsymbol{\alpha}_{j_t}$ 也线性无关知,$t\leqslant r$.从而有,$r=t$. □

定理 5.2 说明一个向量组可能有不同的极大无关组,但极大无关组的所含向量的个数是确定的,由此,我们可以给出向量组秩的概念.

定义 5.10 向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的极大线性无关组中所含向量的个数 r 称为这个**向量组的秩**(rank),记作 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)$.**只含零向量的向量组的秩规定为** 0.

例如,$\boldsymbol{\alpha}_1=\begin{pmatrix}0\\0\end{pmatrix},\boldsymbol{\alpha}_2=\begin{pmatrix}1\\0\end{pmatrix},\boldsymbol{\alpha}_3=\begin{pmatrix}2\\0\end{pmatrix}$ 的极大线性无关组中只有一个向量 $\boldsymbol{\alpha}_2$(或 $\boldsymbol{\alpha}_3$).因此,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 的秩为 1,记作 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=1$.

又如,$\boldsymbol{\alpha}_1=(1,1,1),\boldsymbol{\alpha}_2=(0,1,1),\boldsymbol{\alpha}_3=(0,0,1)$ 是阶梯形向量组,它们线性无关.所以,极大线性无关组就是向量组自身,其秩为 3,记作 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=3$.

有了秩的概念,我们再来看两个向量组之间的关系,有如下结果.

定理 5.3 设 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出,若 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)=r$,$r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)=p$,则 $p\leqslant r$.

证明 设 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 与 $\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_p}$ 分别是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 的极大线性

无关组.

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 和 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 是等价向量组,而 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出,因此,$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 可由 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性表出,那么,$\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_p}$ 就可由 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性表出,又因 $\boldsymbol{\beta}_{j_1},\boldsymbol{\beta}_{j_2},\cdots,\boldsymbol{\beta}_{j_p}$ 线性无关,由定理 5.1 知 $p\leqslant r$. □

例 5.6 已知 $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1$ 线性无关.证明:$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关.

证明 记 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1$.

即 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出,又因为:

$$\boldsymbol{\alpha}_1=\frac{1}{2}(\boldsymbol{\beta}_1+\boldsymbol{\beta}_3-\boldsymbol{\beta}_2),\quad \boldsymbol{\alpha}_2=\frac{1}{2}(\boldsymbol{\beta}_1+\boldsymbol{\beta}_2-\boldsymbol{\beta}_3),\quad \boldsymbol{\alpha}_3=\frac{1}{2}(\boldsymbol{\beta}_2+\boldsymbol{\beta}_3-\boldsymbol{\beta}_1),$$

说明 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 亦可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性表出,它们是等价的向量组,从而

$$r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=r(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1)=3.$$

所以,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关. □

5.3.4 向量(子)空间的基与维数

定义 5.11 设 V 表示 F^n 或 F^n 的子空间,若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是 V 中 s 个线性无关的向量,且 V 中任何向量 $\boldsymbol{\beta}$ 均可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 唯一地线性表示,设为

$$\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_s\boldsymbol{\alpha}_s,$$

则称 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 为**空间 V 的一组基**,$x_1,x_2,\cdots,x_s$ 称为**向量 $\boldsymbol{\beta}$ 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 下的坐标**.s 称为 V 的维数,记为 $\dim V$.

例如,$\boldsymbol{e}_1=(1,0,\cdots,0)^{\mathrm{T}},\boldsymbol{e}_2=(0,1,\cdots,0)^{\mathrm{T}},\cdots,\boldsymbol{e}_n=(0,\cdots,0,1)^{\mathrm{T}}$ 线性无关,F^n 中任何一个向量都可用 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 线性表出,且表达式是唯一的,所以它们是一组基,称为 F^n 中的**自然基**,故 F^n 的维数为 n.自然基是 $\mathbb{R}^3$ 中标准的直角坐标系在 F^n 中的推广,$\mathbb{R}^3$ 中任意三个不共面的向量也可以线性组合出全体向量,称为 $\mathbb{R}^3$ 的仿射坐标系.而 V 中一般的基是 $\mathbb{R}^3$ 中仿射坐标系在 F^n 中的推广.

由基的定义与极大无关组的定义,很容易验证.

定理 5.4 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s\}$ 为 V 的一组基 $\Leftrightarrow\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s\}$ 为 V 中所有向量构成的向量组的极大无关组.

证明略.

由定理 5.4 可知:

(1) “基”与“维数”是特别情况下的“极大无关组”与“秩”,此时的向量组由全体 V 中(无穷多个)向量构成.

(2) 向量空间 V 中的基的选取不唯一,但维数是确定的.

§5.4 矩阵的秩

对于向量组来说,向量组的秩是反映向量组的线性相关性质的不变量.一个矩阵既可以看作由列向量组成的,也可以看作由行向量组成的,那么反映矩阵的这些向量组的线性相关性的不变量是什么呢？本节将引入矩阵的秩的概念,它是反映矩阵的本质的一个不变量.另外可以从理论上讨论清楚线性方程组什么时候有解;当方程组有无穷多解时,又如何把这无穷多个解表达清楚.

5.4.1 矩阵的三种秩

定义 5.12 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵,对 $\boldsymbol{A}$ 按列分块 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$,其中 $\boldsymbol{\alpha}_i(i=1,2,\cdots,n)$ 是 m 维向量,我们称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的秩 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 是**矩阵 $\boldsymbol{A}$ 的列秩**(column rank),记为 $r_c(\boldsymbol{A})$.

对 $\boldsymbol{A}$ 按行分块 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\beta}_1\\\boldsymbol{\beta}_2\\\vdots\\\boldsymbol{\beta}_m\end{pmatrix}$,其中 $\boldsymbol{\beta}_j(j=1,2,\cdots,m)$ 是 n 维行向量,称向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 的秩$r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m)$是**矩阵 $\boldsymbol{A}$ 的行秩**(row rank),记为 $r_r(\boldsymbol{A})$.

例 5.7 对于 $\boldsymbol{A}=\begin{pmatrix}1&2&3\\0&4&5\end{pmatrix}$,$\boldsymbol{A}$ 的列向量组是 $\boldsymbol{\alpha}_1=\begin{pmatrix}1\\0\end{pmatrix}$,$\boldsymbol{\alpha}_2=\begin{pmatrix}2\\4\end{pmatrix}$,$\boldsymbol{\alpha}_3=\begin{pmatrix}3\\5\end{pmatrix}$,显然其秩 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=2$,即 $r_c(\boldsymbol{A})=2$.

$\boldsymbol{A}$ 的行向量组是 $\boldsymbol{\beta}_1=(1,2,3)$,$\boldsymbol{\beta}_2(0,4,5)$,其秩 $r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)=2$,即 $r_r(\boldsymbol{A})=2$.

由上例中看到,$r_r(\boldsymbol{A})=r_c(\boldsymbol{A})$,这个结论在一般情况下是否成立呢？为了讨论这个问题,我们引入矩阵的第三种秩的概念,并且由它作为桥梁,把三种秩统一起来.

定义 5.13 设 $\boldsymbol{A}=(a_{ij})_{m\times n}$是一个 $m\times n$ 的矩阵,任取 k 行 k 列,由位于这些行与列的交叉点上的元素,按原来的次序组成的 k 阶行列式称为 $\boldsymbol{A}$ 中的一个 **k 阶子式**,记作 $\det\left(\boldsymbol{A}_k\begin{pmatrix}i_1&i_2&\cdots&i_k\\j_1&j_2&\cdots&j_k\end{pmatrix}\right)$.

例 $\boldsymbol{A}=\begin{pmatrix}1&2&3&4\\0&0&5&6\\0&0&0&0\end{pmatrix}$.

$$\det\left(A_2\begin{pmatrix}1,2\\1,3\end{pmatrix}\right)=\begin{vmatrix}1&3\\0&5\end{vmatrix}=5,\det\left(A_3\begin{pmatrix}1,2,3\\1,2,4\end{pmatrix}\right)=\begin{vmatrix}1&2&4\\0&0&6\\0&0&0\end{vmatrix}=0.$$

定义 5.14 设 $\boldsymbol{A}$ 是一个 $m\times n$ 矩阵,$\boldsymbol{A}$ 中不为 0 的子式的最高阶数 r,称为**矩阵 $\boldsymbol{A}$ 的秩**,记作 $r(\boldsymbol{A})$,或 $\mathrm{rank}(\boldsymbol{A})$,或秩$(\boldsymbol{A})$.

例如上例,$r\left(\begin{pmatrix}1&2&3&4\\0&0&5&6\\0&0&0&0\end{pmatrix}\right)=2$,因为存在不为 0 的二阶子式 $\det\left(A_2\begin{pmatrix}1,2\\2,3\end{pmatrix}\right)=\begin{vmatrix}2&3\\0&5\end{vmatrix}=10\neq0$,而所有三阶以上的子式均等于 0.

5.4.2 矩阵秩的计算

定理 5.5 初等变换不改变矩阵的秩,即若矩阵 $\boldsymbol{A}$ 可通过初等变换化为矩阵 $\boldsymbol{B}$,则 $r(\boldsymbol{A})=r(\boldsymbol{B})$.

证明 按三种初等变换分别讨论:

(1) 数乘变换 $k\times ⓘ(k\neq0)$:

由于 $k\neq0$,不影响矩阵中任何一个子式的非 0 性;

(2) 对换变换,交换 i,j 两行,$ⓘ\leftrightarrow ⓙ$:

所有子式或不变,或只差一个符号,因而也不影响矩阵中任何一个子式的非 0 性.

(3) 倍加变换,即把第 i 行的 k 倍加到第 j 行.记为 $ⓙ+kⓘ$.

设 $r(\boldsymbol{A})=r$,$\boldsymbol{A}$ 经过一次倍加变换 $ⓙ+kⓘ$ 变为 $\boldsymbol{B}$,设 D 为 $\boldsymbol{B}$ 中任何一个 $r+1$ 阶子式.分为下面三种情况:

① 如果 D 中不包含第 i 行和第 j 行,则 D 也是 $\boldsymbol{A}$ 的 $r+1$ 阶子式,所以 $D=0$;

② 如果 D 同时包含第 i 行和第 j 行,则根据行列式的性质,与原来的 $\boldsymbol{A}$ 的子式相等;

③ 如果 D 包含第 j 行但不包含第 i 行,则 D 可分拆为两个子式之和,其中一个也是 $\boldsymbol{A}$ 的 $r+1$ 阶子式,另一个是 $\boldsymbol{A}$ 的 $r+1$ 阶子式经过对换或倍乘得到,它们都等于 0,所以 $D=0$.

综上,$r(\boldsymbol{B})\leqslant r$,由于倍加变换可逆,$\boldsymbol{B}$ 经过一次倍加变换可变为 $\boldsymbol{A}$.因此得 $r(\boldsymbol{A})\leqslant r(\boldsymbol{B})$,所以 $r(\boldsymbol{B})=r(\boldsymbol{A})=r$,即倍加变换不改变矩阵的秩. □

推论 5.1 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵,$\boldsymbol{P}$ 是 m 阶可逆阵,$\boldsymbol{Q}$ 是 n 阶可逆阵,则

$$r(\boldsymbol{PA})=r(\boldsymbol{AQ})=r(\boldsymbol{PAQ})=r(\boldsymbol{A}).$$

证明 由定理 5.5,定理 4.3 及矩阵乘法的意义可证. □

例 5.8 求矩阵 $\boldsymbol{A}$ 的秩,其中 $\boldsymbol{A}=\begin{pmatrix}1&-1&0&1&1\\2&-2&0&2&2\\1&1&1&0&0\\2&0&1&1&1\end{pmatrix}$.

解 由 $A\to\begin{pmatrix}1&-1&0&1&1\\0&0&0&0&0\\0&2&1&-1&-1\\0&2&1&-1&-1\end{pmatrix}\to\begin{pmatrix}1&-1&0&1&1\\0&2&1&-1&-1\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}=B$，

知 $r(A)=r(B)=2$.

5.4.3 三种秩的统一

我们要证明同一个矩阵的三种秩是相等的.

由前所定义，记矩阵 A 的列秩为 $r_c(A)$，矩阵 A 的行秩为 $r_r(A)$，矩阵 A 的秩为 $r(A)$.

定理 5.6 设 A 是 $m\times n$ 矩阵，则 $r(A)=r_c(A)=r_r(A)$.

证明 设 $r(A)=r$，把 A 用高斯消元法变为阶梯形，由定理 5.5 以及 $r(A)=r$，必然只有 r 行非 0：

$$A=(\alpha_1,\alpha_2,\cdots,\alpha_n)\to\begin{pmatrix}c_{11}&\cdots&c_{1r}&\cdots&c_{1n}\\\vdots&&\vdots&&\vdots\\0&\cdots&c_{rr}&\cdots&c_{rn}\\0&\cdots&0&\cdots&0\\\vdots&&\vdots&&\vdots\\0&\cdots&0&\cdots&0\end{pmatrix},$$

其中 $c_{ii}\neq 0, i=1,\cdots,r$.

c_{ii}所在的列对应的向量 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 构成一组极大无关组，因而 $r(\alpha_1,\alpha_2,\cdots,\alpha_n)=r$，所以 $r(A)=r(\alpha_1,\alpha_2,\cdots,\alpha_n)$，即 $r(A)=r_c(A)$. 把此结论用于 $A^{\mathrm T}$，得 $r(A^{\mathrm T})=r_c(A^{\mathrm T})=r_r(A)$.

又因为 $r(A^{\mathrm T})=r(A)$，所以 $r(A)=r_r(A)$.

综上，$r(A)=r_c(A)=r_r(A)$. □

有了这个定理，我们在分析和运算中，不管是矩阵的秩，还是行秩，或是列秩，它们都一样.

5.4.4 矩阵秩的性质

1° $r(A)=r(A^{\mathrm T})$.

2° $r(A)\leqslant\min\{m,n\}$. 当 $r(A)=m$ 时，称为行满秩. 当 $r(A)=n$ 时，称为列满秩. 当 A 为方阵时，$r(A)=m=n\Leftrightarrow A$ 可逆$\Leftrightarrow\det A\neq 0$.

3° $r(A+B)\leqslant r(A)+r(B)$.

4° $\max\{r(\boldsymbol{A}),r(\boldsymbol{B})\}\leqslant r(\boldsymbol{A}\,\vdots\,\boldsymbol{B})\leqslant r(\boldsymbol{A})+r(\boldsymbol{B})$.

5° $r(\boldsymbol{AB})\leqslant\min\{r(\boldsymbol{A}),r(\boldsymbol{B})\}$.

6° $r\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}=r(\boldsymbol{A})+r(\boldsymbol{B})$.

7° $r\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{C} & \boldsymbol{B}\end{pmatrix}\geqslant r(\boldsymbol{A})+r(\boldsymbol{B})$.

8° $r(\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times p})\geqslant r(\boldsymbol{A})+r(\boldsymbol{B})-n$.

9° 设 $\boldsymbol{A}=\boldsymbol{A}_{m\times n},\boldsymbol{B}=\boldsymbol{B}_{n\times p}$,则有 $\boldsymbol{AB}=\boldsymbol{O}\Rightarrow r(\boldsymbol{A})+r(\boldsymbol{B})\leqslant n$.

以上 9 条性质的部分证明比较有技巧性,留作习题.

小结:矩阵的秩的概念和有关性质的内容比较丰富,而且难度较高,能否掌握是对学习能力的一种考验.

我们小结一下解决这类问题时的思路:

① 考虑矩阵中不为 0 的子式;

② 把矩阵转换为行向量组或列向量组,利用向量组的秩和极大无关组;

③ 利用矩阵的相抵标准形;

④ 利用分块矩阵;

⑤ 灵活应用已得到的秩的性质;

⑥ 用一些最简单的矩阵来记忆和想象.

§5.5 线性方程组的解理论

有了矩阵的秩之后,我们可以对线性方程组问题再展开讨论.关于线性方程组的解的问题,我们似乎在第二章中用高斯消元法解决了,那么还要讨论什么问题呢?首先,高斯消元法只能对具体的每个系数矩阵展开计算,而本节将对抽象的符号矩阵所对应的线性方程组是否有解,给出一个统一标准;其次,在有解的情况下,我们还会将线性方程组的解的结构用 n 维向量空间的向量加以描述.

我们分齐次与非齐次线性方程组的解两部分介绍.

5.5.1 齐次线性方程组有非零解的条件

设齐次线性方程组为

$$A_{m\times n}X=\mathbf{0}, \tag{5.3}$$

其中 $A_{m\times n}\in M_{m,n}(F)$，$F$ 是数域.由于解线性方程组所涉及的运算是线性运算，因此它的解仍是数域 F 上的向量.下面，我们都是在数域 F 上讨论问题.

对齐次线性方程组我们主要讨论它什么时候有非零解，有非零解时解的结构是什么.我们先把前面已经得到的齐次线性方程组有非零解的条件总结如下.

齐次线性方程组有非零解的条件.

(1) 设 A 是方阵，则 $AX=\mathbf{0}$ 有非零解$\Leftrightarrow|A|=0$.

(2) 对矩阵 $A_{m\times n}$，若 $m<n$，则 $A_{m\times n}X=\mathbf{0}$ 有非零解.

我们可以利用前面给出的秩的概念，对齐次线性方程组有非零解的条件给出统一的判定准则.

定理 5.7 齐次线性方程组 $A_{m\times n}X=\mathbf{0}$ 有非零解$\Leftrightarrow r(A_{m\times n})<n$.

证明 虽然我们可以从高斯消元法来证明，但用“秩”这个更先进的武器好一些.

$\Rightarrow$：设 $A_{m\times n}X=\mathbf{0}$ 有非零解 $X=(k_1,k_2,\cdots,k_n)^{\mathrm{T}}\neq\mathbf{0}$，并设

$$A_{m\times n}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n).$$

则有 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\mathbf{0}$，所以 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性相关，因而 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)<n$，所以 $r(A_{m\times n})=r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)<n$.

$\Leftarrow$：设 $r(A_{m\times n})<n$，即 $r(A_{m\times n})=r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)<n$，因而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性相关，于是存在不全为 0 的数 $k_1,k_2,\cdots,k_n$ 使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\mathbf{0}.$$

令

$$X_1=(k_1,k_2,\cdots,k_n)^{\mathrm{T}},$$

则 $A_{m\times n}X_1=\mathbf{0}$ 且 $X_1\neq\mathbf{0}$，所以 $A_{m\times n}X=\mathbf{0}$ 有非零解. □

下面讨论 $A_{m\times n}X=\mathbf{0}$ 在有非零解的情况下，它的解有什么特性.

5.5.2 齐次线性方程组的解空间与基础解系

设齐次线性方程组 $A_{m\times n}X=\mathbf{0}$ 的解集合为 $N(A_{m\times n})$，我们来证明它是 F^n 的子空间.首先，$\mathbf{0}\in N(A_{m\times n})$，所以 $N(A_{m\times n})\neq\varnothing$.其次，还要证明 $N(A_{m\times n})$ 对向量加法和数乘满足封闭性.

(1) 设 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2$ 是 $A_{m\times n}X=\mathbf{0}$ 的解，由于 $A(\boldsymbol{\eta}_1+\boldsymbol{\eta}_2)=A\boldsymbol{\eta}_1+A\boldsymbol{\eta}_2=\mathbf{0}$，所以 $\boldsymbol{\eta}_1+\boldsymbol{\eta}_2\in N(A_{m\times n})$.

(2) 设 $\boldsymbol{\eta}$ 是 $A_{m\times n}X=\mathbf{0}$ 的解，$k\in F$，由于 $A(k\boldsymbol{\eta})=kA\boldsymbol{\eta}=\mathbf{0}$，所以 $k\boldsymbol{\eta}\in N(A_{m\times n})$.

综上，所以 $N(A_{m\times n})$ 是 F^n 的子空间，称为 $A_{m\times n}X=\mathbf{0}$ 的**解空间**，或 $A_{m\times n}$ 的**化零空间**.

如何确定和表示解空间呢？请看下面的讨论.

定义 5.15 设 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_t$ 是 $A_{m\times n}X=\mathbf{0}$ 的一组线性无关解，且 $A_{m\times n}X=\mathbf{0}$ 的任何一个解

均可由 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_t$ 线性表出，则称 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_t$ 为 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的一个**基础解系**.

基础解系就是解空间 $N(\boldsymbol{A}_{m\times n})$ 的一组基.

有了基础解系，我们就可把 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的任何一个解表为

$$\boldsymbol{\eta}=c_1\boldsymbol{\eta}_1+c_2\boldsymbol{\eta}_2+\cdots+c_t\boldsymbol{\eta}_t, \tag{5.4}$$

其中 $c_1,c_2,\cdots,c_t$ 为任意常数.形式(5.4)的解称为 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的**一般解或通解**.

下面我们要讨论基础解系所含向量的个数和如何求此基础解系.

定理 5.8 设 $r(\boldsymbol{A}_{m\times n})=r$，则 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的基础解系中所含向量的个数为 $n-r$.因而 $\dim N(\boldsymbol{A}_{m\times n})=n-r$.

证明 类似于定理 5.7，可以利用矩阵的秩或高斯消元法来证明，但这次由于我们考虑到要同时求出基础解系，所以我们选择利用高斯消元法来证明此定理.首先，用高斯消元法把系数矩阵化为阶梯形，

$$\boldsymbol{A}\to\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1r} & \cdots & c_{1n}\\ & c_{22} & \cdots & c_{2r} & \cdots & c_{2n}\\ & & \ddots & \vdots & & \vdots\\ & & & c_{rr} & \cdots & c_{rn}\\ & & & 0 & \cdots & 0 \end{pmatrix},$$

其中为讨论问题的方便，我们假设 $c_{ii}\neq 0(i=1,2,\cdots,r)$.实际消元的结果不一定是规则的阶梯形，而是不规则的阶梯形，见后面的例子.然后把消成阶梯形的矩阵写回成方程组的形式，并把对应$r+1\sim n$ 列的部分移到右边，得

$$\begin{cases} c_{11}x_1+c_{12}x_2+\cdots+c_{1r}x_r=-c_{1,r+1}x_{r+1}-\cdots-c_{1n}x_n,\\ \qquad c_{22}x_2+\cdots+c_{2r}x_r=-c_{2,r+1}x_{r+1}-\cdots-c_{2n}x_n,\\ \qquad\cdots\cdots\cdots\cdots\\ \qquad\qquad c_{rr}x_r=-c_{r,r+1}x_{r+1}-\cdots-c_{rn}x_n, \end{cases} \tag{5.5}$$

方程组(5.5)与原方程组(5.3)是同解的.

方程组(5.5)中，变量 $x_{r+1},\cdots,x_n$ 可取任意值，称为自由变量.

令 $x_{r+1}=k_1,\cdots,x_n=k_{n-r}$为任意常数，将其代入式(5.5)，则从最后一个方程开始，可逐个解出 $x_r,x_{r-1},\cdots,x_1$.

$$\begin{cases} x_1=b_{11}k_1+\cdots+b_{1,n-r}k_{n-r},\\ x_2=b_{21}k_1+\cdots+b_{2,n-r}k_{n-r},\\ \cdots\cdots\cdots\cdots\\ x_r=b_{r1}k_1+\cdots+b_{r,n-r}k_{n-r},\\ x_{r+1}=k_1,\\ \cdots\cdots\cdots\cdots\\ x_n=k_{n-r}, \end{cases} \tag{5.6}$$

进一步可把式(5.6)的解表为列向量形式：

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = k_1\begin{pmatrix} b_{11} \\ \vdots \\ b_{r1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + k_2\begin{pmatrix} b_{12} \\ \vdots \\ b_{r2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + k_{n-r}\begin{pmatrix} b_{1,n-r} \\ \vdots \\ b_{r,n-r} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{5.7}$$

令

$$\boldsymbol{\eta} = \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix},\quad \boldsymbol{\eta}_1 = \begin{pmatrix} b_{11} \\ \vdots \\ b_{r1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\quad \boldsymbol{\eta}_2 = \begin{pmatrix} b_{12} \\ \vdots \\ b_{r2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix},\quad \cdots,\quad \boldsymbol{\eta}_{n-r} = \begin{pmatrix} b_{1,n-r} \\ \vdots \\ b_{r,n-r} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \tag{5.8}$$

则得

$$\boldsymbol{\eta} = k_1\boldsymbol{\eta}_1 + k_2\boldsymbol{\eta}_2 + \cdots + k_{n-r}\boldsymbol{\eta}_{n-r}, \tag{5.9}$$

其中 $k_1, k_2, \cdots, k_{n-r}$ 为任意数.

下面证明 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}$ 是方程组(5.5)的一个基础解系：

首先证明 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}$ 线性无关,由于

$$(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}) = \begin{pmatrix} \boldsymbol{D} \\ \boldsymbol{I}_{n-r} \end{pmatrix},$$

故 $r(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}) = n-r$,所以 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}$ 线性无关.

其次,式(5.9)说明方程组(5.5)的任何一个解可由 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}$ 线性表出.所以 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}$ 是方程组(5.5)的一个基础解系.

式(5.9)就是(5.3)的通解,式(5.3)的解空间可表为

$$N(\boldsymbol{A}_{m\times n}) = \{k_1\boldsymbol{\eta}_1 + k_2\boldsymbol{\eta}_2 + \cdots + k_{n-r}\boldsymbol{\eta}_{n-r} \mid k_1, k_2, \cdots, k_{n-r} \in F\}.$$

且解空间的维数为 $\dim(N(\boldsymbol{A}_{m\times n})) = n-r$.

但如果我们为了先求出基础解系,可用与前面稍稍不同的做法来解方程(5.5)：对自由变量 $x_{r+1}, \cdots, x_n$ 分别作：

取 $x_{r+1}=1, x_{r+2}=0, \cdots, x_n=0$,代入式(5.5),可逐个解出 $x_r, x_{r-1}, \cdots, x_1$,得到 $\boldsymbol{\eta}_1$.

取 $x_{r+1}=0, x_{r+2}=1, \cdots, x_n=0$,代入式(5.5),可逐个解出 $x_r, x_{r-1}, \cdots, x_1$,得到 $\boldsymbol{\eta}_2$.

……

取 $x_{r+1}=0,x_{r+2}=0,\cdots,x_n=1$，代入式(5.5)，可逐个解出 $x_r,x_{r-1},\cdots,x_1$，得到 $\boldsymbol{\eta}_{n-r}$.
结果与式(5.8)相同. □

从以上证明过程，我们可以总结出，求 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的基础解系和通解的计算步骤：

(1) 用高斯消元法把系数矩阵化成(简化)阶梯形.

(2) 若 $r(\boldsymbol{A})=n$，则方程组只有零解.若 $r(\boldsymbol{A})<n$，则方程组有非零解，做下一步.

(3) 确定自由变量，设为 $x_{r+1},\cdots,x_n$.对自由变量分别取 $(1,0,\cdots,0),(0,1,\cdots,0),\cdots,(0,0,\cdots,0,1)$ 解出其他变量，求得基础解系 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-r}$.

(4) 写出 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的通解：$\boldsymbol{\eta}=k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2+\cdots+k_{n-r}\boldsymbol{\eta}_{n-r}$，其中 $k_1,\cdots,k_{n-r}$ 为任意常数.

下面举例说明.

例 5.9　求以下齐次线性方程组的通解

$$\begin{cases} x_1+3x_2-5x_3-x_4+2x_5=0, \\ 2x_1+6x_2-8x_3+5x_4+3x_5=0, \\ x_1+3x_2-3x_3+6x_4+x_5=0. \end{cases} \tag{5.10}$$

解　(1) 用高斯消元法把系数矩阵化为阶梯形：

$$\begin{pmatrix} 1 & 3 & -5 & -1 & 2 \\ 2 & 6 & -8 & 5 & 3 \\ 1 & 3 & -3 & 6 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -5 & -1 & 2 \\ 0 & 0 & 2 & 7 & -1 \\ 0 & 0 & 2 & 7 & -1 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 3 & -5 & -1 & 2 \\ 0 & 0 & 2 & 7 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 0 & 33/2 & -1/2 \\ 0 & 0 & 1 & 7/2 & -1/2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{5.11}$$

(2) 由于 $r(\boldsymbol{A})=2<n=5$，故方程组有非零解，且基础解系所含向量的个数为 3.

(3) 由(5.11)可看出，自由变量为 x_2,x_4,x_5，对自由变量分别取

令 $x_2=1,x_4=0,x_5=0$，可得 $x_1=-3,x_3=0$，得 $\boldsymbol{\eta}_1=(-3,1,0,0,0)^{\mathrm{T}}$；

令 $x_2=0,x_4=1,x_5=0$，可得 $x_1=-\dfrac{33}{2},x_3=-\dfrac{7}{2}$，得 $\boldsymbol{\eta}_2=\left(-\dfrac{33}{2},0,-\dfrac{7}{2},1,0\right)^{\mathrm{T}}$；

令 $x_2=0,x_4=0,x_5=1$，可得 $x_1=\dfrac{1}{2},x_3=\dfrac{1}{2}$，得 $\boldsymbol{\eta}_3=\left(\dfrac{1}{2},0,\dfrac{1}{2},0,1\right)^{\mathrm{T}}$.

(4) 得到基础解系为

$$\boldsymbol{\eta}_1=(-3,1,0,0,0)^{\mathrm{T}},\quad \boldsymbol{\eta}_2=\left(-\frac{33}{2},0,-\frac{7}{2},1,0\right)^{\mathrm{T}},\quad \boldsymbol{\eta}_3=\left(\frac{1}{2},0,\frac{1}{2},0,1\right)^{\mathrm{T}}.$$

通解为 $\boldsymbol{\beta}=c_1\boldsymbol{\eta}_1+c_2\boldsymbol{\eta}_2+c_3\boldsymbol{\eta}_3$，其中 c_1,c_2,c_3 为任意数.

通解和基础解系也可用下列方式得到：

令 $x_2=k_1,x_4=k_2,x_5=k_3$ 为任意数，则可求得

$$\begin{cases} x_1 = -3k_1 - \dfrac{33}{2}k_2 + \dfrac{1}{2}k_3, \\ x_2 = k_1, \\ x_3 = -\dfrac{7}{2}k_2 + \dfrac{1}{2}k_3, \\ x_4 = k_2, \\ x_5 = k_3. \end{cases} \tag{5.12}$$

(注意:为了下一步写成向量形式,上下按任意数对齐.)

把解式(5.12)写成向量形式:

$$\boldsymbol{\beta} = k_1\begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + k_2\begin{pmatrix} -\dfrac{33}{2} \\ 0 \\ -\dfrac{7}{2} \\ 1 \\ 0 \end{pmatrix} + k_3\begin{pmatrix} \dfrac{1}{2} \\ 0 \\ \dfrac{1}{2} \\ 0 \\ 1 \end{pmatrix},$$

其中 k_1,k_2,k_3 为任意数.这就是齐次方程组(5.10)的通解,式(5.10)的基础解系为

$$\boldsymbol{\eta}_1 = \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\eta}_2 = \begin{pmatrix} -\dfrac{33}{2} \\ 0 \\ -\dfrac{7}{2} \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\eta}_3 = \begin{pmatrix} \dfrac{1}{2} \\ 0 \\ \dfrac{1}{2} \\ 0 \\ 1 \end{pmatrix}.$$

结果是相同的.

思考题:能否取 x_3,x_4,x_5 为自由变量?为什么?能否取 x_2,x_3,x_4 为自由变量?

5.5.3 非齐次线性方程组有解的条件

下面主要讨论非齐次线性方程组的两个问题:一是用矩阵的秩来表达有解的条件;二是在有解的情况下,它的解的结构.

设非齐次线性方程组为

$$\boldsymbol{A}_{m\times n}\boldsymbol{X} = \boldsymbol{b}, \tag{5.13}$$

$$\boldsymbol{A}_{m\times n} \in M_{m,n}(F).$$

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 为 $\boldsymbol{A}$ 的列向量.则(5.13)等价于

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{b}.$$

于是(5.13)有解⇔向量 $\boldsymbol{b}$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表出

$$\Leftrightarrow r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{b})\Leftrightarrow r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{b}).$$

进一步,当(5.13)有解时,若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关,即 $r(\boldsymbol{A})=n$ ($\boldsymbol{A}$ 列满秩).由 5.2 节的性质(10)知,$\boldsymbol{b}$ 可被 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 唯一地线性表出,即(5.13)有唯一解.反之,若(5.13)有唯一解,即 $\boldsymbol{b}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 唯一地线性表出,说明必有 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关,即 $r(\boldsymbol{A})=n$.反证:否则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性相关,则存在不全为零的 $k_1,k_2,\cdots,k_n$,使得

$$\boldsymbol{0}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n.$$

设 $\boldsymbol{b}=l_1\boldsymbol{\alpha}_1+l_2\boldsymbol{\alpha}_2+\cdots+l_n\boldsymbol{\alpha}_n$,上述两式相加得

$$\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{0}=(k_1+l_1)\boldsymbol{\alpha}_1+(k_2+l_2)\boldsymbol{\alpha}_2+\cdots+(k_n+l_n)\boldsymbol{\alpha}_n.$$

这与 $\boldsymbol{b}$ 的唯一表出性矛盾.

将上述总结为定理:

定理 5.9 非齐次线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 有解⇔$r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{b})$.

进一步,$r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{b})=n\Leftrightarrow\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一解.

$r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{b})<n\Leftrightarrow\boldsymbol{Ax}=\boldsymbol{b}$ 有非唯一的解.

注 由下一节的讨论知,当 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解且不唯一时,必有无穷多解.

下面讨论非齐次线性方程组有解时解的结构.

5.5.4 非齐次线性方程组有解时解的结构

首先讨论非齐次线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的解与对应的齐次线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$(又称为导出组)的解的关系.

(1) 设 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2$ 都是 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的解,则 $\boldsymbol{\xi}_1-\boldsymbol{\xi}_2$ 是对应的齐次线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的解.

(2) 设 $\boldsymbol{\xi}_0$ 是 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的解,$\boldsymbol{\eta}$ 是对应的 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的解,则 $\boldsymbol{\eta}+\boldsymbol{\xi}_0$ 仍是 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的解.

(3) 设 $\boldsymbol{\xi}_0$ 是 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的某个特解,则 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的任一解 $\boldsymbol{\gamma}$ 可表示为

$$\boldsymbol{\gamma}=\boldsymbol{\xi}_0+\boldsymbol{\eta},$$

其中 $\boldsymbol{\eta}$ 是对应的 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的解.

第(3)条的证明请看下面定理 5.10 的证明.根据第(3)条,当 $\boldsymbol{\eta}$ 取遍 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的所有解时,$\boldsymbol{\gamma}=\boldsymbol{\xi}_0+\boldsymbol{\eta}$ 就取遍 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的所有解.

定理 5.10 设非齐次线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 满足 $r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{b})=r<n$,$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-r}$ 是导出组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的基础解系,$\boldsymbol{\xi}_0$ 是 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的一个特解,则 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的通解为

$$\boldsymbol{\xi}=\boldsymbol{\xi}_0+k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2+\cdots+k_{n-r}\boldsymbol{\eta}_{n-r},\tag{5.14}$$

其中 $k_1,k_2,\cdots,k_{n-r}$ 为任意数.

证明 要证明(5.14)是(5.13)的通解,需要证明两点:一是(5.14)是(5.13)的解. 二是(5.13)的任何一个解均为(5.14)中 $k_1,k_2,\cdots,k_{n-r}$ 取定的某组值.

第一点只要把 $\boldsymbol{\xi}=\boldsymbol{\xi}_0+k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2+\cdots+k_{n-r}\boldsymbol{\eta}_{n-r}$ 代入(5.13)进行验算:

$$\boldsymbol{A\xi}=\boldsymbol{A\xi}_0+k_1\boldsymbol{A\eta}_1+\cdots+k_{n-r}\boldsymbol{A\eta}_{n-r}=\boldsymbol{b}.$$

证明第二点:设 $\boldsymbol{\beta}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$ 是(5.13)的任一个解. 由于 $\boldsymbol{\beta}-\boldsymbol{\xi}_0$ 是齐次线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{0}$ 的解,根据齐次线性方程组的基础解系的性质,必有 $k_1,k_2,\cdots,k_{n-r}\in F$ 使

$$\boldsymbol{\beta}-\boldsymbol{\xi}_0=k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2+\cdots+k_{n-r}\boldsymbol{\eta}_{n-r},$$

即

$$\boldsymbol{\beta}=\boldsymbol{\xi}_0+k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2+\cdots+k_{n-r}\boldsymbol{\eta}_{n-r},$$

所以(5.14)是 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的通解.

由此得(5.13)的解集合为

$$N(\boldsymbol{A},\boldsymbol{b})=\{\boldsymbol{\xi}_0+k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2+\cdots+k_{n-r}\boldsymbol{\eta}_{n-r}\mid k_1,k_2,\cdots,k_{n-r}\in F\},$$

但它不是 F^n 的子空间.(请思考为什么?) □

求解非齐次线性方程组的计算步骤:

(1) 写出方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{X}=\boldsymbol{b}$ 的增广矩阵:$(\boldsymbol{A},\boldsymbol{b})$.

(2) 对 $(\boldsymbol{A},\boldsymbol{b})$ 作高斯消元法,使其变为阶梯形.

(3) 分以下几种情况讨论:我们用秩的语言改写为

① 若 $r(\boldsymbol{A})<r(\boldsymbol{A},\boldsymbol{b})$,则方程组无解.

② 若 $r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{b})=r$,又分以下两种情况:

- 若 $r=n$,则方程组有唯一解.
- 若 $r<n$,则对应的齐次线性方程组中含有 $n-r$ 个自由变量,确定自由变量,设为 $x_{r+1},\cdots,x_n$. 对自由变量分别取 $(1,0,\cdots,0),(0,1,\cdots,0),\cdots,(0,\cdots,0,1)$ 解出其他变量,求得基础解系 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-r}$.

(4) 写出消元后的非齐次线性方程组,令自由变量都取 0,求出非齐次线性方程组的一个特解 $\boldsymbol{\xi}_0$.

(5) 写出非齐次线性方程组的通解 $\boldsymbol{\xi}=\boldsymbol{\xi}_0+k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2+\cdots+k_{n-r}\boldsymbol{\eta}_{n-r}$,其中 $k_1,k_2,\cdots,k_{n-r}$ 为任意数.

下面我们通过例题来说明方法.

例 5.10 解方程组

$$\begin{cases}x_1-x_2-x_3+3x_5=-1,\\2x_1-2x_2-x_3+2x_4+4x_5=-2,\\3x_1-3x_2-x_3+4x_4+5x_5=-3,\\x_1-x_2+x_3+x_4+8x_5=2,\end{cases}$$

并将其解表示为通解形式.

解 ① 对增广矩阵进行初等行变换,化为简化的阶梯形:

$$(\boldsymbol{A},\boldsymbol{b})=\left(\begin{array}{ccccc:c}1&-1&-1&0&3&-1\\2&-2&-1&2&4&-2\\3&-3&-1&4&5&-3\\1&-1&1&1&8&2\end{array}\right)\to\left(\begin{array}{ccccc:c}1&-1&0&0&7&1\\0&0&1&0&4&2\\0&0&0&1&-3&-1\\0&0&0&0&0&0\end{array}\right),\qquad(5.15)$$

得到 $r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{b})=3$.方程组有解.

② 由非主元素所在的列,确定自由变量为 x_2,x_5,并令 $(x_2,x_5)=(1,0)$ 与 $(0,1)$.得到对应齐次线性方程组(导出组)的基础解系为(部分信息包含在(5.15)的第2、5列)

$$\boldsymbol{\eta}_1=(1,1,0,0,0)^{\mathrm{T}},$$

$$\boldsymbol{\eta}_2=(-7,0,-4,3,1)^{\mathrm{T}}.$$

③ 令 $(x_2,x_5)=(0,0)$.由简化阶梯形(5.15)的最后一列知,非齐次的特解为

$$\boldsymbol{\xi}_0=(1,0,2,-1,0)^{\mathrm{T}}.$$

④ 定出非齐次方程组的通解

$$\boldsymbol{\xi}=\boldsymbol{\xi}_0+c_1\boldsymbol{\eta}_1+c_2\boldsymbol{\eta}_2=(1+c_1-7c_2,c_1,2-4c_2,-1+3c_2,c_2)^{\mathrm{T}}.$$

其中 $c_1,c_2\in F$ 可取任意值.

例 5.11 解方程组

$$\begin{cases}x_1+x_2+x_3+x_4+x_5=1,\\3x_1+2x_2+x_3+x_4-3x_5=0,\\5x_1+4x_2+3x_3+3x_4-x_5=2.\end{cases}$$

解 用高斯消元法将 $(\boldsymbol{A},\boldsymbol{b})$ 变化为简化阶梯形:

$$(\boldsymbol{A},\boldsymbol{b})=\left(\begin{array}{ccccc:c}1&1&1&1&1&1\\3&2&1&1&-3&0\\5&4&3&3&-1&2\end{array}\right)\to\left(\begin{array}{ccccc:c}1&0&-1&-1&-5&-2\\0&1&2&2&6&3\\0&0&0&0&0&0\end{array}\right),\qquad(5.16)$$

实际上通解的所有信息都已包含在简化阶梯形(5.16)之中.$r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{b})=2$,方程组有解,自由变量为 x_3,x_4,x_5.

令 $(x_3,x_4,x_5)=(0,0,0)$,得特解为 $\boldsymbol{\xi}_0=(-2,3,0,0,0)^{\mathrm{T}}$.

令 $(x_3,x_4,x_5)=(1,0,0),(0,1,0)$ 与 $(0,0,1)$,得导出组基础解系为

$$\boldsymbol{\eta}_1=(1,-2,1,0,0)^{\mathrm{T}},$$

$$\boldsymbol{\eta}_2=(1,-2,0,1,0)^{\mathrm{T}},$$

$$\boldsymbol{\eta}_3=(5,-6,0,0,1)^{\mathrm{T}}.$$

故原方程组解的通解为

$$\boldsymbol{\xi}=\boldsymbol{\xi}_0+k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2+k_3\boldsymbol{\eta}_3,$$

其中 $k_1,k_2,k_3\in F$ 可取任意值.

5.5.5 $\mathbb{R}^3$中三个平面的位置关系

$\mathbb{R}^3$中一个关于 x,y,z 的三元一次(线性)方程对应一个平面. 下面我们来讨论一下$\mathbb{R}^3$中三个平面的位置关系. 从几何上看,三个平面可能有以下八种关系(见图 5.1).

从代数上看,这相当于考虑下面的线性方程组

$$\begin{cases} a_1x+b_1y+c_1z=d_1, \\ a_2x+b_2y+c_2z=d_2, \\ a_3x+b_3y+c_3z=d_3 \end{cases} \quad (a_i,b_i,c_i,d_i\in\mathbb{R},i=1,2,3) \tag{5.17}$$

的解的情况.

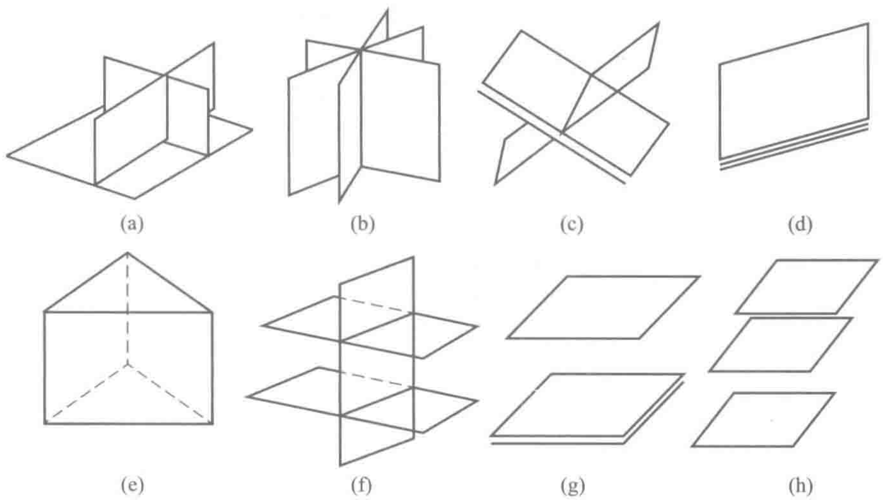


图 5.1 $\mathbb{R}^3$中三个平面可能的位置关系

令

$$\boldsymbol{A}=\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix},\qquad \boldsymbol{d}=\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix},\qquad \begin{array}{l} \boldsymbol{\beta}_1^{\mathrm{T}}=(a_1,b_1,c_1), \\ \boldsymbol{\beta}_2^{\mathrm{T}}=(a_2,b_2,c_2), \\ \boldsymbol{\beta}_3^{\mathrm{T}}=(a_3,b_3,c_3), \end{array}$$

$$\boldsymbol{\gamma}_1^{\mathrm{T}}=(a_1,b_1,c_1,d_1),$$
$$\boldsymbol{\gamma}_2^{\mathrm{T}}=(a_2,b_2,c_2,d_2),$$
$$\boldsymbol{\gamma}_3^{\mathrm{T}}=(a_3,b_3,c_3,d_3),$$

即系数矩阵 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\beta}_1^{\mathrm{T}}\\\boldsymbol{\beta}_2^{\mathrm{T}}\\\boldsymbol{\beta}_3^{\mathrm{T}}\end{pmatrix}$，增广矩阵 $(\boldsymbol{A},\boldsymbol{d})=\begin{pmatrix}\boldsymbol{\gamma}_1^{\mathrm{T}}\\\boldsymbol{\gamma}_2^{\mathrm{T}}\\\boldsymbol{\gamma}_3^{\mathrm{T}}\end{pmatrix}$.

利用线性方程组解的理论，以及线性相关理论，我们得到以上八种几何情形的代数等价条件．结果如表 5.1：

表 5.1　$\mathbb{R}^3$ 中三个平面位置关系的代数等价条件

方程组解数	解集合(交集)	等价代数条件(i)(ii)		几何情形
有唯一解	一点	(i) $r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{d})=3$		(1)
有无穷多解	一直线	(i) $r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{d})=2$	(ii) $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3$ 两两线性无关	(2)
			(ii) $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3$ 中有两个线性相关	(3)
	一平面	(i) $r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{d})=1$		(4)
无解	空集 $\varnothing$	(i) $r(\boldsymbol{A})=2$, $r(\boldsymbol{A},\boldsymbol{d})=3$	(ii) $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 两两线性无关	(5)
			(ii) $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 中有两个线性相关	(6)
		(i) $r(\boldsymbol{A})=1$, $r(\boldsymbol{A},\boldsymbol{d})=2$	(ii) $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3$ 中有两个线性相关	(7)
			(ii) $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3$ 两两线性无关	(8)

§5.6　欧氏空间

在三维空间 $\mathbb{R}^3$ 中，我们曾定义了向量的内积(点乘)．现在要把内积概念推广到实向量空间中去，并把定义了内积的实向量空间称为欧氏空间(Euclid space).

5.6.1　内积与欧氏空间

在三维几何空间中，建立坐标系后就可定义向量的长度和两个向量的夹角，然后用向量的长度和两个向量的夹角来定义两个向量的内积．但在线性空间里我们还不知道什么是向量的长度和两个向量的夹角，因而无法用它们来定义两个向量的内积．我们可以从内积的性质来定义两个向量的内积．

定义 5.16　设 $V=V(\mathbb{R}^n)$ 是实向量空间，$\varphi: V\times V\to\mathbb{R}$ 是一个映射，$\forall\,\boldsymbol{\alpha},\boldsymbol{\beta}\in V$，记 $\varphi(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{\beta})\in\mathbb{R}$，并满足：$\forall\,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in V$，有：

(1) $(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\beta},\boldsymbol{\alpha})$　(对称性)；

(2) $(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})$　(分配律)；

(3) $\forall\, c\in\mathbb{R}$ 有 $(c\boldsymbol{\alpha},\boldsymbol{\beta})=c(\boldsymbol{\alpha},\boldsymbol{\beta})$　(常数可提出)；

(4) $(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$ 且 $(\boldsymbol{\alpha},\boldsymbol{\alpha})=0\Leftrightarrow\boldsymbol{\alpha}=\mathbf{0}$ （正定性），

则称 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 是 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的**内积**，并把定义了内积的实向量空间称为**欧几里得空间**，简称**欧氏空间**.

强调一点：以前在介绍向量空间的时候，数域可以是其他数域，但欧氏空间的数域必须是实数域，欧氏空间的维数不一定是有限的．但本书研究的仅是有限维的欧氏空间．

下面通过一些实际例子来说明内积的概念．

例 5.12 在 $\mathbb{R}^n$ 中，对任意两个向量 $\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 和 $\boldsymbol{\beta}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$，定义

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}=\sum_{i=1}^{n}x_iy_i, \tag{5.18}$$

则容易验证其满足定义 5.16，所以是 $\mathbb{R}^n$ 中的内积．此内积称为标准内积．

在同一个线性空间中我们可以定义不同的内积，请看下例．

例 5.13 设 $\boldsymbol{A}=\begin{pmatrix}1&2\\2&5\end{pmatrix}$，在 $\mathbb{R}^2$ 中，对任意两个向量 $\boldsymbol{\alpha}=(x_1,x_2)^{\mathrm{T}}$ 和 $\boldsymbol{\beta}=(y_1,y_2)^{\mathrm{T}}$，定义

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta}=x_1y_1+2x_1y_2+2x_2y_1+5x_2y_2. \tag{5.19}$$

下面我们来验证式(5.19)满足定义 5.16：

(1) $(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta}$，$(\boldsymbol{\beta},\boldsymbol{\alpha})=\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}$，两边求转置：$(\boldsymbol{\beta},\boldsymbol{\alpha})=(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha})^{\mathrm{T}}=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\beta}=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta}=(\boldsymbol{\alpha},\boldsymbol{\beta})$.（因为 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$.）

(2) $(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha}+\boldsymbol{\beta})^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\gamma}=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\gamma}+\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\gamma}=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})$.

(3) $(c\boldsymbol{\alpha},\boldsymbol{\beta})=c(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta})=c(\boldsymbol{\alpha},\boldsymbol{\beta})$．类似有 $(\boldsymbol{\alpha},c\boldsymbol{\beta})=c(\boldsymbol{\alpha},\boldsymbol{\beta})$（其中 $c\in\mathbb{R}$）.

(4) $(\boldsymbol{\alpha},\boldsymbol{\alpha})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=x_1^2+4x_1x_2+5x_2^2=(x_1+2x_2)^2+x_2^2\geqslant 0$，

$(\boldsymbol{\alpha},\boldsymbol{\alpha})=\mathbf{0}\Leftrightarrow(x_1+2x_2)^2+x_2^2=0\Leftrightarrow x_1=x_2=0\Leftrightarrow\boldsymbol{\alpha}=\mathbf{0}$.

综上，式(5.19)是 $\mathbb{R}^2$ 中的内积．

5.6.2 内积与度量

定义向量 $\boldsymbol{\alpha}$ 的长度如下：

$$|\boldsymbol{\alpha}|=\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})}.$$

所以，$|c\boldsymbol{\alpha}|=|c|\,|\boldsymbol{\alpha}|$，$c\in\mathbb{R}$．长度为 1 的向量称为**单位向量**.对于任意向量 $\boldsymbol{\alpha}\neq\mathbf{0}$，$\dfrac{1}{|\boldsymbol{\alpha}|}\boldsymbol{\alpha}$ 为单位向量，这称为对 $\boldsymbol{\alpha}$ 的单位化.

那么如何计算两向量的夹角呢？为此，我们还是回到三维几何空间中的内积的定义：

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=|\boldsymbol{\alpha}|\,|\boldsymbol{\beta}|\cos<\boldsymbol{\alpha},\boldsymbol{\beta}>$$

因此对夹角有

$$\cos<\boldsymbol{\alpha},\boldsymbol{\beta}>=\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{|\boldsymbol{\alpha}|\,|\boldsymbol{\beta}|}. \tag{5.20}$$

并规定 $0\leqslant<\boldsymbol{\alpha},\boldsymbol{\beta}>\leqslant\pi$.

但是要使式(5.20)对任何欧氏空间有意义,必须 $|(\boldsymbol{\alpha},\boldsymbol{\beta})|\leqslant|\boldsymbol{\alpha}||\boldsymbol{\beta}|$. 幸运的是下面的定理保证了此条件成立.

定理 5.11(柯西-施瓦茨(Cauchy-Schwarz)不等式) 设 V 是欧氏空间,则 $\forall\,\boldsymbol{\alpha},\boldsymbol{\beta}\in V$ 有

$$(\boldsymbol{\alpha},\boldsymbol{\beta})^2\leqslant|\boldsymbol{\alpha}|^2|\boldsymbol{\beta}|^2,$$

且等式成立当且仅当 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 线性相关.

证明 任取实数 x,考虑内积 $(\boldsymbol{\alpha}+x\boldsymbol{\beta},\boldsymbol{\alpha}+x\boldsymbol{\beta})$:

$$(\boldsymbol{\alpha}+x\boldsymbol{\beta},\boldsymbol{\alpha}+x\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{\alpha})+2x(\boldsymbol{\alpha},\boldsymbol{\beta})+x^2(\boldsymbol{\beta},\boldsymbol{\beta}).$$

由定义 5.16 的(4),当 $\forall\,x\in\mathbb{R}$ 且 $\boldsymbol{\alpha}+x\boldsymbol{\beta}\neq\mathbf{0}$ 时有 $(\boldsymbol{\alpha}+x\boldsymbol{\beta},\boldsymbol{\alpha}+x\boldsymbol{\beta})>0$,即

$$(\boldsymbol{\alpha},\boldsymbol{\alpha})+2x(\boldsymbol{\alpha},\boldsymbol{\beta})+x^2(\boldsymbol{\beta},\boldsymbol{\beta})>0,\tag{5.21}$$

由于 x 是任意实数,式(5.21)左边的二次式无实根,因而由二次代数方程的根的性质得

$$4(\boldsymbol{\alpha},\boldsymbol{\beta})^2-4(\boldsymbol{\alpha},\boldsymbol{\alpha})(\boldsymbol{\beta},\boldsymbol{\beta})<0,$$

所以

$$(\boldsymbol{\alpha},\boldsymbol{\beta})^2<|\boldsymbol{\alpha}|^2|\boldsymbol{\beta}|^2.$$

当 $\exists\,x\in\mathbb{R}$ 使 $\boldsymbol{\alpha}+x\boldsymbol{\beta}=\mathbf{0}$ 时,有 $\boldsymbol{\alpha}+x\boldsymbol{\beta}=\mathbf{0}\Leftrightarrow\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 线性相关. 且这时有 $\boldsymbol{\alpha}=-x\boldsymbol{\beta}$,于是 $(\boldsymbol{\alpha},\boldsymbol{\beta})^2=x^2(\boldsymbol{\beta},\boldsymbol{\beta})^2=|\boldsymbol{\alpha}|^2|\boldsymbol{\beta}|^2$.

综上,得 $(\boldsymbol{\alpha},\boldsymbol{\beta})^2\leqslant|\boldsymbol{\alpha}|^2|\boldsymbol{\beta}|^2$. □

注 向量的长度和两向量的夹角依赖内积的定义. 例如,在 $\mathbb{R}^2$ 中,对两个向量 $\boldsymbol{\alpha}=(1,2)^{\mathrm{T}},\boldsymbol{\beta}=(-1,1)^{\mathrm{T}}$.

按标准内积(5.18),得 $|\boldsymbol{\alpha}|=\sqrt{5}$, $|\boldsymbol{\beta}|=\sqrt{2}$, $\cos<\boldsymbol{\alpha},\boldsymbol{\beta}>=\dfrac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{|\boldsymbol{\alpha}||\boldsymbol{\beta}|}=\dfrac{1}{\sqrt{10}}$.

按内积定义(5.19),得 $|\boldsymbol{\alpha}|=\sqrt{29}$, $|\boldsymbol{\beta}|=\sqrt{2}$, $\cos<\boldsymbol{\alpha},\boldsymbol{\beta}>=\dfrac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{|\boldsymbol{\alpha}||\boldsymbol{\beta}|}=\dfrac{7}{\sqrt{58}}$.

由两向量的夹角的定义,我们得到:若 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$,则 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 垂直或正交;

若 $(\boldsymbol{\alpha},\boldsymbol{\beta})>0$,则 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 夹锐角;若 $(\boldsymbol{\alpha},\boldsymbol{\beta})<0$,则 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 夹钝角.

当 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 共线(平行)时,有 $\cos<\boldsymbol{\alpha},\boldsymbol{\beta}>=\pm1$ 或 $|(\boldsymbol{\alpha},\boldsymbol{\beta})|=|\boldsymbol{\alpha}||\boldsymbol{\beta}|$.

正如前面指出的,零向量与任何向量都正交,也可看做与任何向量都平行.

以下无特别声明时,欧氏空间 $\mathbb{R}^n$ 中的内积都指的是标准内积,因而向量的长度和两向量的夹角都是确定的.

5.6.3 一个矩阵秩的例子

利用齐次线性方程组的解理论与欧氏空间中的内积性质,我们可证明如下关于矩阵秩的性质. 这个性质很重要,在后文中将会用到.

例 5.14 设 $\boldsymbol{A}\in M_{m\times n}(\mathbb{R})$.证明 $r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=r(\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}})$.

证明 考虑如下两个齐次线性方程组

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0} \text{ 与 } \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}.$$

下面,我们来证明它们是同解方程组,即 $N(\boldsymbol{A})=N(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$.

一方面,$\forall\,\boldsymbol{\alpha}\in N(\boldsymbol{A})$,有 $\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{0}\Rightarrow\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{0}=\boldsymbol{0}\Rightarrow\boldsymbol{\alpha}\in N(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$.

另一方面,$\forall\,\boldsymbol{\alpha}\in N(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$,有 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{0}$.两边左乘 $\boldsymbol{\alpha}^{\mathrm{T}}$,得

$$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=(\boldsymbol{A}\boldsymbol{\alpha},\boldsymbol{A}\boldsymbol{\alpha})=|\boldsymbol{A}\boldsymbol{\alpha}|^2=\boldsymbol{\alpha}^{\mathrm{T}}\cdot\boldsymbol{0}=0.$$

故有 $$\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{0},\text{即 }\boldsymbol{\alpha}\in N(\boldsymbol{A}).$$

综上,$N(\boldsymbol{A})=N(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$.从而有

$$\dim N(\boldsymbol{A})=n-r(\boldsymbol{A})=\dim N(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=n-r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$$

$$\Rightarrow r(\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}).$$ □

注 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 为 $n\times n$ 的实对称方阵,由上结论知,$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 可逆$\Leftrightarrow r(\boldsymbol{A})=n\Leftrightarrow\boldsymbol{A}$ 列满秩.

5.6.4 标准正交基和正交矩阵

一个不含零向量的向量组如果其中的向量两两正交,就称为一个**正交向量组**.它有以下性质.

定理 5.12 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 为欧氏空间中的一个正交向量组,则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关.

证明 证明方法是常规的,按线性无关定义来做.

设 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_s\boldsymbol{\alpha}_s=\boldsymbol{0}$,用 $\boldsymbol{\alpha}_i$ 对等式两边做内积,由正交性得

$$x_i(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)=0,$$

由于$(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)>0$,故 $x_i=0,i=1,2,\cdots,s$.

所以 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关. □

如果欧氏空间中的一组基是一个正交向量组,则称这组基为**正交基**.如果一组正交基的每个向量都是单位向量,则称这组基为**标准正交基**或**单位正交基**.例如,在$\mathbb{R}^n$中选定标准内积,则自然基 $\boldsymbol{e}_1=(1,0,\cdots,0)^{\mathrm{T}},\boldsymbol{e}_2=(0,1,\cdots,0)^{\mathrm{T}},\cdots,\boldsymbol{e}_n=(0,\cdots,0,1)^{\mathrm{T}}$ 就是一组标准正交基.在$\mathbb{R}^3$中,$\boldsymbol{\varepsilon}_1=\left(\frac{1}{\sqrt3},\frac{1}{\sqrt3},\frac{1}{\sqrt3}\right)^{\mathrm{T}},\boldsymbol{\varepsilon}_2=\left(\frac{1}{\sqrt2},0,-\frac{1}{\sqrt2}\right)^{\mathrm{T}},\boldsymbol{\varepsilon}_3=\left(\frac{1}{\sqrt6},-\frac{2}{\sqrt6},\frac{1}{\sqrt6}\right)^{\mathrm{T}}$ 是一组标准正交基.

为讨论标准正交基的性质,我们要引进正交矩阵的概念.

定义 5.17 设 $\boldsymbol{A}\in M_n(\mathbb{R})$,若满足 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{I}$,则称 $\boldsymbol{A}$ 是**正交矩阵**.

例如,单位矩阵是正交矩阵.

正交矩阵有以下性质:

(1) $|\boldsymbol{A}|=1$ 或 $|\boldsymbol{A}|=-1$;

（2）正交矩阵 $\boldsymbol{A}$ 可逆，且有 $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{T}}$；

（3）正交矩阵 $\boldsymbol{A}$ 满足：$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$；

（4）$\boldsymbol{A}$ 是正交矩阵，则 $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{T}}$ 也是正交矩阵；

（5）$\boldsymbol{A},\boldsymbol{B}$ 是正交矩阵，则 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 都是正交矩阵.

不难证明，正交矩阵和标准正交基有如下关系：$\boldsymbol{A}$ 是正交矩阵的充要条件是 $\boldsymbol{A}$ 的列向量组是$\mathbb{R}^n$的一组标准正交基.

5.6.5 施密特(Schmidt)正交化方法

现在研究如何把欧氏空间一组线性无关的向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s(s\leqslant n)$ 改造成单位正交的向量组 $\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_s$，并且要求 $\boldsymbol{\gamma}_i(i=1,2,\cdots,s)$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_i$ 的线性组合的方法，即所谓**施密特正交化方法**.

第一步用正交向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 代替 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$；使 $\boldsymbol{\beta}_i(i=1,\cdots,n)$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_i$ 的线性组合.

第二步把 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 单位化为 $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_n$.

先来看二维空间的情况.

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是一组基，我们来求 $L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)$ 的正交基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$. 按要求 $\boldsymbol{\beta}_1$ 是 $\boldsymbol{\alpha}_1$ 的线性组合，$\boldsymbol{\beta}_2$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 的线性组合，为方便起见，可设

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1,\ \boldsymbol{\beta}_2=\boldsymbol{\alpha}_2+k\boldsymbol{\beta}_1.$$

根据 $\boldsymbol{\beta}_2\perp\boldsymbol{\beta}_1$，应有

$$(\boldsymbol{\beta}_2,\boldsymbol{\beta}_1)=(\boldsymbol{\alpha}_2+k\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)=(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)+k(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)=0.$$

由于 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1\neq\boldsymbol{0}$，知 $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)\neq\boldsymbol{0}$，解出

$$k=-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)},$$

从而

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1,$$

这样的 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是 $L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)$ 的正交基（注：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是等价的向量组）. 再单位化，令

$$\boldsymbol{\gamma}_1=\frac{1}{|\boldsymbol{\beta}_1|}\boldsymbol{\beta}_1,\ \boldsymbol{\gamma}_2=\frac{1}{|\boldsymbol{\beta}_2|}\boldsymbol{\beta}_2.$$

那么 $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2$ 就是合乎要求的单位正交向量组.

在三维空间中，如 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关，用上述方法作出 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 之后，对 $\boldsymbol{\beta}_3$，可令 $\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3+l\boldsymbol{\beta}_1+m\boldsymbol{\beta}_2$.

由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 两两正交，得

$$(\boldsymbol{\beta}_3,\boldsymbol{\beta}_1)=(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)+l(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)=0,$$
$$(\boldsymbol{\beta}_3,\boldsymbol{\beta}_2)=(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)+m(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)=0,$$

因此，$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\dfrac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\dfrac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2$.

归纳地，我们有下面的定理.

定理 5.13 n 维欧氏空间 V 中，任意 $s(s\leqslant n)$ 个线性无关的向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$，可用施密特正交化方法转化成一个正交向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$，其中

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1,\ \boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1,\cdots,$$

$$\boldsymbol{\beta}_s=\boldsymbol{\alpha}_s-\frac{(\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_s,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2-\cdots-\frac{(\boldsymbol{\alpha}_s,\boldsymbol{\beta}_{s-1})}{(\boldsymbol{\beta}_{s-1},\boldsymbol{\beta}_{s-1})}\boldsymbol{\beta}_{s-1},$$

而且 $L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)=L(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s)$，进而令

$$\boldsymbol{\gamma}_i=\frac{\boldsymbol{\beta}_i}{|\boldsymbol{\beta}_i|},\ i=1,2,\cdots,s.$$

可得单位正交向量组 $\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_s$.

例 5.15 用施密特正交化方法把 $\mathbb{R}^4$ 的如下向量组化为一组单位正交的向量组.

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\1\\1\\0\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}-1\\0\\-1\\0\end{pmatrix},\quad \boldsymbol{\alpha}_3=\begin{pmatrix}-1\\2\\3\\1\end{pmatrix}.$$

解 先正交化，有

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1=\begin{pmatrix}1\\1\\1\\0\end{pmatrix},$$

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1=\begin{pmatrix}-1\\0\\-1\\0\end{pmatrix}+\frac{2}{3}\begin{pmatrix}1\\1\\1\\0\end{pmatrix}=\frac{1}{3}\begin{pmatrix}-1\\2\\-1\\0\end{pmatrix},$$

$$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2$$

$$=\begin{pmatrix}-1\\2\\3\\1\end{pmatrix}-\frac{4}{3}\begin{pmatrix}1\\1\\1\\0\end{pmatrix}-\frac{1}{3}\begin{pmatrix}-1\\2\\-1\\0\end{pmatrix}=\begin{pmatrix}-2\\0\\2\\1\end{pmatrix}.$$

再单位化,有

$$\boldsymbol{\gamma}_1=\frac{\boldsymbol{\beta}_1}{|\boldsymbol{\beta}_1|}=\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\\0\end{pmatrix},$$

$$\boldsymbol{\gamma}_2=\frac{\boldsymbol{\beta}_2}{|\boldsymbol{\beta}_2|}=\frac{1}{\sqrt{6}}\begin{pmatrix}-1\\2\\-1\\0\end{pmatrix},$$

$$\boldsymbol{\gamma}_3=\frac{\boldsymbol{\beta}_3}{|\boldsymbol{\beta}_3|}=\frac{1}{3}\begin{pmatrix}-2\\0\\2\\1\end{pmatrix}.$$

$\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3$ 就是一组单位正交的向量组.

施密特正交化方法的公式有些复杂,但是有规律,初学者需记住公式,多做多练. 下一章我们将运用施密特正交化的方法,来讨论实对称阵的对角化问题.

5.6.6　可逆实矩阵的 QR 分解

将上述施密特正交化的过程用矩阵乘积的形式表示出来,即得到如下关于可逆实矩阵的 QR 分解的结论.

定理 5.14　若 $\boldsymbol{A}$ 为可逆实矩阵,则存在正交阵 $\boldsymbol{Q}$ 与主对角元为正数的上三角阵 $\boldsymbol{R}$,使得

$$\boldsymbol{A}=\boldsymbol{QR}.$$

证明　设 n 阶可逆实矩阵 $\boldsymbol{A}$ 的列向量为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$,则 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 构成 $\mathbb{R}^n$ 的一组基. 由 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 可通过施密特正交化方法得到 $\mathbb{R}^n$ 的一组单位正交基. 首先,做正交化,得一组两两正交的非零向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$,即

$$\begin{cases}\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1,\\ \boldsymbol{\beta}_2=k_{12}\boldsymbol{\beta}_1+\boldsymbol{\alpha}_2,\\ \cdots\cdots\cdots\cdots\\ \boldsymbol{\beta}_n=k_{1n}\boldsymbol{\beta}_1+k_{2n}\boldsymbol{\beta}_2+\cdots+k_{n-1,n}\boldsymbol{\beta}_{n-1}+\boldsymbol{\alpha}_n,\end{cases}\tag{5.22}$$

其中 $k_{ij}=-\dfrac{(\boldsymbol{\alpha}_j,\boldsymbol{\beta}_i)}{(\boldsymbol{\beta}_i,\boldsymbol{\beta}_i)}\quad(1\leqslant i<j,j=2,3,\cdots,n)$.

由(5.22)式可反解出 $\boldsymbol{\alpha}_i$,即

$$\begin{cases}\boldsymbol{\alpha}_1=\boldsymbol{\beta}_1,\\ \boldsymbol{\alpha}_2=-k_{12}\boldsymbol{\beta}_1+\boldsymbol{\beta}_2,\\ \boldsymbol{\alpha}_n=-k_{1n}\boldsymbol{\beta}_1-\cdots-k_{n-1,n}\boldsymbol{\beta}_{n-1}+\boldsymbol{\beta}_n.\end{cases}\tag{5.23}$$

将(5.23)式表示为矩阵等式,为

$$\boldsymbol{A}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)\begin{pmatrix}1 & -k_{12} & \cdots & -k_{1n}\\ & & \ddots & \vdots\\ & & \ddots & -k_{n-1,n}\\ & & & 1\end{pmatrix}.$$

进一步,再单位化,表示为矩阵乘法得

$$\boldsymbol{A}=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n)\begin{pmatrix}\frac{1}{|\boldsymbol{\beta}_1|} & & \\ & \ddots & \\ & & \frac{1}{|\boldsymbol{\beta}_n|}\end{pmatrix}\begin{pmatrix}|\boldsymbol{\beta}_1| & & \\ & \ddots & \\ & & |\boldsymbol{\beta}_n|\end{pmatrix}\begin{pmatrix}1 & -k_{12} & \cdots & -k_{1n}\\ & & \ddots & \vdots\\ & & \ddots & -k_{n-1,n}\\ & & & 1\end{pmatrix}$$

$$=\left(\frac{\boldsymbol{\beta}_1}{|\boldsymbol{\beta}_1|},\cdots,\frac{\boldsymbol{\beta}_n}{|\boldsymbol{\beta}_n|}\right)\begin{pmatrix}|\boldsymbol{\beta}_1| & * & \cdots & *\\ & |\boldsymbol{\beta}_2| & \ddots & \vdots\\ \boldsymbol{O} & & \ddots & *\\ & & & |\boldsymbol{\beta}_n|\end{pmatrix}\triangleq \boldsymbol{QR}.$$

上式中的 $\boldsymbol{Q}$ 由单位正交基组成,故为正交矩阵,而上三角阵 $\boldsymbol{R}$ 的主对角元 $|\boldsymbol{\beta}_i|>0(i=1,2,\cdots,n)$. □

注 QR 分解只适用于可逆实矩阵,它将施密特正交化的公式用矩阵乘积的方式更加直观明了地表示出来.不仅如此,在一般的理论推导中,QR 分解也能发挥作用.

习　题　五

1. 判断下列向量组的线性相关性：

(1) $\boldsymbol{\alpha}_1=(1,-1,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(2,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_3=(3,1,2)^{\mathrm{T}}$;

(2) $\boldsymbol{\beta}_1=(1,-1,1,2)^{\mathrm{T}},\boldsymbol{\beta}_2=(-1,1,1,-1)^{\mathrm{T}},\boldsymbol{\beta}_3=(2,-2,0,3)^{\mathrm{T}},\boldsymbol{\beta}_4=(1,0,1,3)^{\mathrm{T}}$.

2. 当 k 满足什么条件时，向量组 $\boldsymbol{\alpha}_1=(1,0,0,k)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,k,0,0)^{\mathrm{T}},\boldsymbol{\alpha}_3=(0,1,1,0)^{\mathrm{T}}$，$\boldsymbol{\alpha}_4=(0,0,k,1)^{\mathrm{T}}$线性无关？

3. 已知4个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性相关，且其中任意3个向量都线性无关，求证：必有4个非零的数 k_1,k_2,k_3,k_4，使得 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\alpha}_3+k_4\boldsymbol{\alpha}_4=\mathbf{0}$ 成立.

4. 已知向量 $\boldsymbol{\alpha}_1=(\lambda,\lambda,\lambda)^{\mathrm{T}},\boldsymbol{\alpha}_2=(\lambda,2\lambda-1,\lambda)^{\mathrm{T}},\boldsymbol{\alpha}_3=(2,3,\lambda+3)^{\mathrm{T}},\boldsymbol{\beta}=(1,1,2\lambda-1)^{\mathrm{T}}$，当 λ 为何值时，$\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出，且表达式不唯一？

5. 求下列向量组的秩，并求出一个极大线性无关组.

(1) $\boldsymbol{\alpha}_1=(1,0,1,0,1)^{\mathrm{T}},\boldsymbol{\alpha}_2=(0,1,0,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_3=(2,1,2,1,2)^{\mathrm{T}},\boldsymbol{\alpha}_4=(2,1,0,1,2)^{\mathrm{T}}$;

(2) $\boldsymbol{\beta}_1=(1,1,1,-1)^{\mathrm{T}},\boldsymbol{\beta}_2=(1,1,-1,1)^{\mathrm{T}},\boldsymbol{\beta}_3=(1,2,1,1)^{\mathrm{T}}$.

6. 已知向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$，若有另一向量组 $\boldsymbol{\beta}_1=2\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta}_3=-\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2$，求证：$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性相关.

7. 证明第5.4.4节中矩阵秩的性质1°—9°.

8. 求下列齐次线性方程组的基础解系.

$$\begin{cases} x_1+x_2+2x_3-x_4=0,\\ 2x_1+x_2+x_3-x_4=0,\\ 2x_1+2x_2+x_3+2x_4=0. \end{cases}$$

9. 设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶矩阵，且 $\boldsymbol{AB}=\boldsymbol{O}$，求证：$r(\boldsymbol{A})+r(\boldsymbol{B})\leqslant n$.

10. 求下列矩阵的秩：

(1) $\boldsymbol{A}=\begin{pmatrix}1&2&4&1\\1&3&1&5\\2&0&2&2\\1&-1&-2&1\end{pmatrix}$;　(2) $\boldsymbol{B}=\begin{pmatrix}1&1&2&1&2\\2&0&1&-1&5\\0&2&3&3&-1\\1&1&0&-1&4\end{pmatrix}$.

11. 已知矩阵 $\boldsymbol{A}=\begin{pmatrix}1&2&-1\\3&-1&0\\2&x&1\end{pmatrix}$，$\boldsymbol{B}$ 是3阶非零矩阵，若 $\boldsymbol{AB}=\boldsymbol{O}$，求 $r(\boldsymbol{A})$ 和 $r(\boldsymbol{B})$.

12. 已知非齐次线性方程组

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = -1, \\ 4x_1 + 3x_2 + 5x_3 - x_4 = -1, \\ ax_1 + x_2 + 3x_3 + bx_4 = 1 \end{cases}$$

有 3 个线性无关的解.

(1) 求证:方程组系数矩阵 $\boldsymbol{A}$ 的秩 $r(\boldsymbol{A})=2$.

(2) 求 a,b 的值及方程组的通解.

13. 求证:在欧氏空间 $\mathbb{R}^n$ 中,若 $\boldsymbol{\beta}$ 与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 都正交,则 $\boldsymbol{\beta}$ 与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的任一线性组合也正交.

第六章 矩阵的特征值问题

在实际问题中，常常会遇到共振现象，如火车经过一座桥梁时，若火车振动的频率与桥梁的固有频率相同时，桥梁会剧烈震荡．一般来说，一个力学系统有它自己的固有频率，当外力的频率与固有频率相同时将会引起共振.这在设计桥梁时非常重要，一定要使桥梁的固有频率远离火车的振动频率.类似的问题也出现在声学系统、电路系统等，在数学上，可用矩阵的特征值来反映这类问题.

在数学领域，也有一些问题，需要涉及矩阵的特征值问题.例如，微分方程组的求解，多个随机变量的转移过程等.

在第八章中，我们还将用特征值理论来分析空间$\mathbb{R}^3$中的二次曲面及其分类问题.

§6.1 矩阵的特征值与特征向量

6.1.1 基本概念

首先我们给出方阵的特征值与特征向量的概念，因为特别重要，用定义形式给出.

定义 6.1 设 $\boldsymbol{A}$ 是数域 F 上的 n 阶方阵，若存在数 $\lambda(\lambda \in F)$ 和非零向量 $\boldsymbol{X}(\boldsymbol{X} \in F^n)$ 使得

$$\boldsymbol{A}\boldsymbol{X} = \lambda \boldsymbol{X}, \tag{6.1}$$

则称 λ 是 $\boldsymbol{A}$ 的**特征值**(eigenvalue)，$\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的属于特征值 λ 的**特征向量**(eigenvector).

为了有点感性认识，先看一些特殊的例子：零方阵 $\boldsymbol{O}$ 的特征值为 0，所有非零向量都满足式(6.1)，所以非零向量都是特征向量.单位矩阵 $\boldsymbol{I}$ 的特征值是 1，所以非零向量也都满足式(6.1)，所有非零向量都是单位矩阵 $\boldsymbol{I}$ 的特征向量.这两个例子太特殊了，我们还是来看一个一般的例子.

例 6.1 求下列矩阵 $\boldsymbol{A}$ 的所有特征值和特征向量：

$$\boldsymbol{A}=\begin{pmatrix}1&-2\\1&4\end{pmatrix}.$$

解 设 λ 是 $\boldsymbol{A}$ 的特征值，且属于特征值 λ 的特征向量为 $\boldsymbol{X}=(x_1,x_2)^{\mathrm{T}}$，则由 $\boldsymbol{AX}=\lambda\boldsymbol{X}$ 得

$$\begin{pmatrix}1&-2\\1&4\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\lambda\begin{pmatrix}x_1\\x_2\end{pmatrix},$$

移项得

$$\begin{pmatrix}\lambda-1&2\\-1&\lambda-4\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\mathbf{0}. \tag{6.2}$$

根据齐次线性方程组有非零解的条件，得

$$\begin{vmatrix}\lambda-1&2\\-1&\lambda-4\end{vmatrix}=0,$$

即 $\lambda^2-5\lambda+6=0$，解得 $\lambda_1=2,\lambda_2=3$. 这就是 $\boldsymbol{A}$ 的两个特征值.

下面我们对每个特征值求对应的特征向量：

对 $\lambda_1=2$，代入式(6.2)中，得到一个齐次线性方程组

$$\begin{pmatrix}1&2\\-1&-2\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\mathbf{0}.$$

得它的解为 $\boldsymbol{\eta}_1=k_1\begin{pmatrix}2\\-1\end{pmatrix}$，$k_1$ 为任意数.

由于特征向量必须为非零向量，所以属于 $\lambda_1=2$ 的特征值的特征向量为

$$\boldsymbol{X}_1=k_1\begin{pmatrix}2\\-1\end{pmatrix}\quad(k_1\neq 0),$$

我们通常只考虑线性无关的特征向量，因此得到属于 $\lambda_1=2$ 的线性无关的特征向量为

$$\boldsymbol{X}_1=\begin{pmatrix}2\\-1\end{pmatrix}.$$

对于 $\lambda_2=3$，用类似的方法求出对应的特征向量为

$$\boldsymbol{X}_2=\begin{pmatrix}1\\-1\end{pmatrix}.$$

6.1.2 特征值与特征向量的计算方法

由上一小节的例子，可以总结出以下的计算方法.

设 $\boldsymbol{A}=(a_{ij})_{n\times n}$，满足 $\boldsymbol{AX}=\lambda\boldsymbol{X}$.

（1）求特征值的方法：解代数方程

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=0.$$

此方程称为 $\boldsymbol{A}$ 的**特征方程**，详细写出来为

$$\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}=0, \tag{6.3}$$

它是关于 λ 的 n 次代数方程，它的根不一定是实数，但在复数范围内必有 n 个根（代数基本定理），所以在复数范围内任何方阵必有 n 个特征值，但其中可能有重根. 如果特征值是复数，那么特征向量自然也是复数向量，因此，特征值问题一般应在空间 $\mathbb{C}^n$ 中讨论. 例如矩阵 $\boldsymbol{A}=\begin{pmatrix}1 & -2\\ 3 & 2\end{pmatrix}$，特征方程 $\begin{vmatrix}\lambda-1 & 2\\ -3 & \lambda-2\end{vmatrix}=0$ 即 $\lambda^2-3\lambda+8=0$，特征值是复数：$\lambda_{1,2}=\dfrac{3\pm\sqrt{23}\,\mathrm{i}}{2}$.

（2）求特征向量的方法：

假设式（6.3）有 s 个不同的根：$\lambda_1,\lambda_2,\cdots,\lambda_s$. 对每一个特征值 λ_i 解下列相应的齐次线性方程组：

$$(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}. \tag{6.4}$$

由于方程组（6.4）的系数行列式为 0，必有非零解，求出式（6.4）的基础解系就是 $\boldsymbol{A}$ 的属于特征值 λ_i 的全部线性无关的特征向量. 通常我们只要求求出全部线性无关的特征向量就可以了.

再看一个特殊的例子. 对以下的对角阵

$$\boldsymbol{A}=\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}, \quad a_1,a_2,\cdots,a_n \text{ 为互不相同的系数.}$$

如何求它的特征值与特征向量呢？按上述方法：

（1）计算方程 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=0$ 的根，显然得到全部特征值为

$$\lambda_1=a_1,\quad \lambda_2=a_2,\quad \cdots,\quad \lambda_n=a_n.$$

（2）对每一个特征值 $\lambda_i=a_i$ 计算特征向量：由式（6.4）得方程组

$$\begin{pmatrix} a_i - a_1 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & a_i - a_n \end{pmatrix} \boldsymbol{X} = \boldsymbol{0},$$

它的基础解系为 $\boldsymbol{X}_i = \boldsymbol{e}_i = (0,\cdots,1,\cdots,0)^{\mathrm{T}}, i=1,2,\cdots,n$.

所以全部线性无关的特征向量为自然基的各基向量.

易证特征值有下列运算规律:

(i) 设 λ_0 是 $\boldsymbol{A}$ 的特征值,则 $k\lambda_0$ 是 $k\boldsymbol{A}$ 的特征值,λ_0^h 是 $\boldsymbol{A}^h$(h 是非负整数)的特征值.

(ii) 设 λ_0 是 $\boldsymbol{A}$ 的特征值,$f(x)$是任一多项式,则 $f(\lambda_0)$是 $f(\boldsymbol{A})$的特征值.

6.1.3 特征多项式及其性质

在求特征值时,要解一个代数方程——特征方程(6.3),即 $|\lambda \boldsymbol{I}-\boldsymbol{A}|=0$,它的左边是一个关于 λ 的 n 次多项式,令

$$f_{\boldsymbol{A}}(\lambda) = |\lambda \boldsymbol{I} - \boldsymbol{A}|,$$

则称它是 $\boldsymbol{A}$ 的**特征多项式**.

设 $f_{\boldsymbol{A}}(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$,则 $f_{\boldsymbol{A}}(\lambda)$有以下性质:

(1) $f_{\boldsymbol{A}^{\mathrm{T}}}(\lambda)=f_{\boldsymbol{A}}(\lambda)$,即 $\boldsymbol{A}$ 与 $\boldsymbol{A}^{\mathrm{T}}$ 有相同的特征值 λ.

(2) 特征多项式的部分系数与矩阵 $\boldsymbol{A}$ 的元素的关系.

由

$$f_{\boldsymbol{A}}(\lambda) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

$$= \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0. \tag{6.5}$$

易见它的两个特殊的系数 a_0 和 a_{n-1}与 $\boldsymbol{A}$ 的元素有以下关系:

$$a_0 = (-1)^n |\boldsymbol{A}|, \quad a_{n-1} = -\sum_{i=1}^{n} a_{ii}. \tag{6.6}$$

记 $\operatorname{tr} \boldsymbol{A} = \sum_{i=1}^{n} a_{ii}$,称为**方阵 $\boldsymbol{A}$ 的迹**,则 $a_{n-1}=-\operatorname{tr} \boldsymbol{A}$.

(3) 特征多项式的部分系数与 $\boldsymbol{A}$ 的特征值的关系.

由代数基本定理,$f_{\boldsymbol{A}}(\lambda)$在复数范围内有 n 个根,设 $f_{\boldsymbol{A}}(\lambda)$的全部根为

$$\lambda_1, \quad \lambda_2, \quad \cdots, \quad \lambda_n,$$

则 $f_A(\lambda)$ 可表为

$$f_A(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n).$$

因而得到 $f_A(\lambda)$ 的系数有另一组关系：

$$a_0=(-1)^n\prod_{i=1}^{n}\lambda_i,\quad a_{n-1}=-\sum_{i=1}^{n}\lambda_i. \tag{6.7}$$

把式(6.6)(6.7)写在一起，若 $f_A(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$，则

$$a_0=(-1)^n\prod_{i=1}^{n}\lambda_i=(-1)^n|\boldsymbol{A}|, \tag{6.8}$$

$$a_{n-1}=-\sum_{i=1}^{n}\lambda_i=-\sum_{i=1}^{n}a_{ii}=-\operatorname{tr}\boldsymbol{A}, \tag{6.9}$$

因而有以下结果.

(4) 矩阵的秩与 0 特征值的关系：由式(6.8)得到

$\boldsymbol{A}$ 的所有特征值都不为 $0\Leftrightarrow\boldsymbol{A}$ 可逆（0 是 $\boldsymbol{A}$ 的一个特征值 $\Leftrightarrow|\boldsymbol{A}|=0$）.

6.1.4 特征向量的性质与特征子空间

设 λ_0 是 $\boldsymbol{A}$ 的特征值，且 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 是 $\boldsymbol{A}$ 的属于 λ_0 的特征向量，则我们得到以下结果：

(1) $k\boldsymbol{X}_1(k\neq0)$，$\boldsymbol{X}_1+\boldsymbol{X}_2(\neq\boldsymbol{0}$ 时）也是 $\boldsymbol{A}$ 的属于 λ_0 的特征向量.

(2) 令

$$V_{\lambda_0}=\{\boldsymbol{X}\in\mathbb{C}^n\mid\boldsymbol{A}\boldsymbol{X}=\lambda_0\boldsymbol{X}\}, \tag{6.10}$$

则 V_{λ_0} 是 $\mathbb{C}^n$ 的子空间，称为 $\boldsymbol{A}$ 的属于 λ_0 的**特征子空间**.注意，特征子空间除了属于 λ_0 的全体特征向量外，还含有零向量，式(6.10)也反映出了这一点.

(3) 设 $f_A(\lambda)$ 有 s 个不同的根：$\lambda_1,\lambda_2,\cdots,\lambda_s$.令 V_{λ_i} 是 $\boldsymbol{A}$ 的属于 λ_i 的特征子空间，$m_i=\dim(V_{\lambda_i})$ 为特征子空间 V_{λ_i} 的维数，则 m_i 是属于 λ_i 的线性无关的特征向量的个数 $\dim N(\lambda_i\boldsymbol{I}-\boldsymbol{A})=n-r(\lambda_i\boldsymbol{I}-\boldsymbol{A})$.

(4) 属于不同特征值的特征向量线性无关.

我们只对两个特征值的情况证明此性质，读者不难自己证明多个特征值且每个特征值有多个线性无关的特征向量的情况.设 λ_1 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{X}_1$ 是其对应的一个特征向量.设 λ_2 也是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{X}_2$ 是其对应的一个特征向量，且 $\lambda_1\neq\lambda_2$.要证 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 线性无关.设

$$k_1\boldsymbol{X}_1+k_2\boldsymbol{X}_2=\boldsymbol{0}, \tag{6.11}$$

左乘 $\boldsymbol{A}$，得

$$k_1\boldsymbol{A}\boldsymbol{X}_1+k_2\boldsymbol{A}\boldsymbol{X}_2=k_1\lambda_1\boldsymbol{X}_1+k_2\lambda_2\boldsymbol{X}_2=\boldsymbol{0}. \tag{6.12}$$

把式(6.11)$\times\lambda_1-$(6.12)，得 $k_2(\lambda_1-\lambda_2)\boldsymbol{X}_2=\boldsymbol{0}$，因为 $\lambda_1-\lambda_2\neq0$ 和 $\boldsymbol{X}_2\neq\boldsymbol{0}$，故得 $k_2=0$，代入式(6.11)，得 $k_1=0$.所以 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 线性无关.

下面再看一个例子.

例 6.2 求如下矩阵 $\boldsymbol{A}$ 的所有特征值和特征向量.

$$\boldsymbol{A}=\begin{pmatrix}2 & 2 & -1\\ -1 & 0 & 1\\ -1 & 0 & 2\end{pmatrix}.$$

解 根据前面给出的计算方法,分以下两步:

(1) 求特征值.计算特征多项式,并求它的根.

$$f_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2 & -2 & 1\\ 1 & \lambda & -1\\ 1 & 0 & \lambda-2\end{vmatrix}$$

$$=\begin{vmatrix}0 & -2 & -(\lambda-1)(\lambda-3)\\ 0 & \lambda & -(\lambda-1)\\ 1 & 0 & \lambda-2\end{vmatrix}=(\lambda-1)\begin{vmatrix}-2 & -\lambda+3\\ \lambda & -1\end{vmatrix}$$

$$=(\lambda-1)(\lambda^2-3\lambda+2)$$

$$=(\lambda-1)^2(\lambda-2),$$

得到特征值为 $\lambda_1=1$(2 重根),$\lambda_2=2$.

(2) 对每一个特征值求线性无关的特征向量.解齐次线性方程组

$$\begin{pmatrix}\lambda-2 & -2 & 1\\ 1 & \lambda & -1\\ 1 & 0 & \lambda-2\end{pmatrix}\boldsymbol{X}=\boldsymbol{0}.$$

对 $\lambda_1=1$,作高斯消元法:

$$\begin{pmatrix}-1 & -2 & 1\\ 1 & 1 & -1\\ 1 & 0 & -1\end{pmatrix}\to\begin{pmatrix}-1 & -2 & 1\\ 0 & -1 & 0\\ 0 & -2 & 0\end{pmatrix}\to\begin{pmatrix}1 & 2 & -1\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix},$$

由基础解系得到属于 $\lambda_1=1$ 的线性无关的特征向量为

$$\boldsymbol{X}_1=\begin{pmatrix}1\\0\\1\end{pmatrix}.$$

类似可得属于 $\lambda_2=2$ 的线性无关的特征向量为

$$\boldsymbol{X}_2=\begin{pmatrix}0\\1\\2\end{pmatrix}.$$

§6.2 矩阵的相似与对角化

首先看一个例子，设方阵 $A\in M_n(F)$.若 A 有 n 个互不相同的特征值 $\lambda_1,\cdots,\lambda_n\in F$ 对任意 $1\leqslant i\leqslant n$,必有特征向量 X_i,使得 $AX_i=\lambda_iX_i(1\leqslant i\leqslant n)$.

将上述 n 个列向量排起来得到

$$(AX_1,\cdots,AX_n)=(\lambda X_1,\cdots,\lambda_nX_n),$$

即

$$A(X_1,\cdots,X_n)=(X_1,\cdots,X_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}.$$

令 $P=(X_1,\cdots,X_n)$, $\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$为一个 n 阶方阵.由上节属于不同特征值的特征向量的线性无关性知，P 为可逆方阵，故有

$$AP=P\Lambda\Rightarrow P^{-1}AP=\Lambda.$$

上式说明通过可逆矩阵 P,我们可把原方阵 A 变为对角阵 Λ,其中 Λ 包含了 A 的特征值信息，P 包含了 A 的特征向量的信息.下面我们将讨论一般的情况.我们首先给出矩阵相似的概念，然后讨论矩阵相似问题与矩阵特征值问题的联系，之后给出矩阵与对角阵相似的条件与方法.最后，再看一个可对角化矩阵的简单应用.

6.2.1 矩阵的相似

定义 6.2 设有 n 阶方阵 A 与 B,若有可逆阵 P 使得 $P^{-1}AP=B$,则称 A 与 B 相似，记作 $A\sim B$.

相似的矩阵有很多不变量和相同的性质.

(1) 若 $A\sim B$,则 $A^m\sim B^m$,因而有 $f(A)\sim f(B)$,其中 $f(x)$为任意一个多项式.这是因为 $A\sim B\Rightarrow$有可逆阵 P 使 $P^{-1}AP=B\Rightarrow(P^{-1}AP)^m=B^m\Rightarrow P^{-1}A^mP=B^m\Rightarrow A^m\sim B^m$.设 $f(x)=x^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$,则

$$\begin{aligned}P^{-1}f(A)P&=P^{-1}A^mP+a_{m-1}P^{-1}A^{m-1}P+\cdots+a_1P^{-1}AP+a_0I,\\&=B^m+a_{m-1}B^{m-1}+\cdots+a_1B+a_0I=f(B).\end{aligned}$$

所以 $f(A)\sim f(B)$.

(2) 相似的矩阵有相同的特征多项式,因而有相同的特征值.设 $A\sim B\Rightarrow$ 有可逆阵 P 使 $P^{-1}AP=B$,因为

$$f_A(\lambda)=|\lambda I-A|=|P^{-1}||\lambda I-A||P|=|\lambda I-P^{-1}AP|$$
$$=|\lambda I-B|=f_B(\lambda),$$

所以 A 与 B 有相同的特征值.

下面我们来看特征向量的关系.

设 λ_0 为 A 的特征值,X 为对应的特征向量,则

$$AX=\lambda_0X\Rightarrow P^{-1}(AX)=\lambda_0P^{-1}X$$
$$\Rightarrow P^{-1}APP^{-1}X=\lambda_0P^{-1}X$$
$$\Rightarrow BP^{-1}X=\lambda_0P^{-1}X,$$

所以 $P^{-1}X$ 是 B 的属于 λ_0 的特征向量.

(3) 相似的矩阵有相同的行列式.

(4) 相似的矩阵有相同的迹.

但是把性质(2)反过来,两个矩阵有相同的特征值,不一定相似,所以如何判断两个矩阵相似的问题是一个困难的问题.

6.2.2 矩阵的对角化问题

让我们回忆一下矩阵的初等变换,一个矩阵经过一系列初等变换后可化为一个非常简单的矩阵,我们也可把 $P^{-1}AP=B$ 称为 A 用 P 作相似变换变为 B.那么一个方阵能与一个什么样的简单的方阵相似呢?最简单的方阵莫过于对角阵了,所以我们的问题就是一个方阵能否与一个对角阵相似?如果一个方阵与一个对角阵相似,我们就说此方阵可对角化.本节就是讨论方阵可对角化的条件.它与矩阵的特征值与特征向量有密切关系,有以下结果.

矩阵可对角化的充要条件:

定理 6.1 n 阶方阵 A 可对角化的充要条件是 A 有 n 个线性无关的特征向量.

证明 先证必要性.设 n 阶方阵 A 可对角化,则有可逆阵 P 使

$$P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n),\tag{6.13}$$

令 $P=(X_1,X_2,\cdots,X_n)$,$\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$,则式(6.13)可写成

$$AP=P\Lambda$$
$$AP=(AX_1,AX_2,\cdots,AX_n),$$
$$P\Lambda=(X_1,X_2,\cdots,X_n)\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}$$

$$= (\lambda_1 X_1, \lambda_2 X_2, \cdots, \lambda_n X_n).$$

故得

$$AX_k = \lambda_k X_k, \quad k = 1,2,\cdots,n.$$

所以 $X_k, k=1,2,\cdots,n$,都是 A 的特征向量,又因 P 可逆,它们线性无关.

再证充分性.设 A 有 n 个线性无关的特征向量 $X_1, X_2, \cdots, X_n$.设它们对应的特征值分别为 $\lambda_1, \lambda_2, \cdots, \lambda_n$.令 $P=(X_1, X_2, \cdots, X_n)$,则

$$AP = (AX_1, AX_2, \cdots, AX_n) = (\lambda_1 X_1, \lambda_2 X_2, \cdots, \lambda_n X_n)$$

$$= (X_1, X_2, \cdots, X_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = P\Lambda,$$

因而 $P^{-1}AP=\Lambda$,所以 A 可对角化. □

由定理 6.1 可得以下推论:

推论 6.1 若 n 阶方阵 A 有 n 个不同的特征值,则 A 可对角化.

证明 由于每个特征值至少有一个特征向量,属于不同的特征值的特征向量是线性无关的,因而 A 有 n 个线性无关的特征向量,由定理 6.1 知 A 可对角化. □

下面来看一些例子.

例 6.3 判断下列矩阵能否对角化,如能对角化,求出 P 使 $P^{-1}AP=\Lambda$(对角阵).

(1) $A=\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$; (2) $A=\begin{pmatrix} 2 & 2 & -1 \\ -1 & 0 & 1 \\ -1 & 0 & 2 \end{pmatrix}$.

解 (1) 先求 A 的特征值.特征多项式为

$$f_A(\lambda) = \begin{vmatrix} \lambda-2 & -1 & -1 \\ -1 & \lambda-2 & -1 \\ -1 & -1 & \lambda-2 \end{vmatrix} = \begin{vmatrix} 0 & -\lambda+1 & \lambda^2-4\lambda+3 \\ 0 & \lambda-1 & -\lambda+1 \\ -1 & -1 & \lambda-2 \end{vmatrix}$$

$$= -(\lambda-1)^2 \begin{vmatrix} -1 & \lambda-3 \\ 1 & -1 \end{vmatrix} = (\lambda-1)^2(\lambda-4),$$

所以得到 A 的特征值为

$$\lambda_1 = 1(\text{重数为 } 2), \quad \lambda_2 = 4.$$

再求 A 的特征向量.

对特征值 $\lambda_1=1$,解齐次方程组 $(\lambda_1 I-A)X=\mathbf{0}$:

$$(\lambda_1 I - A) = \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

得线性无关的两个特征向量为

$$X_{11}=(-1,0,1)^{\mathrm{T}},\quad X_{12}=(-1,1,0)^{\mathrm{T}},$$

对特征值 $\lambda_2=4$,解齐次方程组 $(\lambda_2 I-A)X=0$:

$$(\lambda_2 I-A)=\begin{pmatrix}2&-1&-1\\-1&2&-1\\-1&-1&2\end{pmatrix}\to\begin{pmatrix}1&1&-2\\0&3&-3\\0&-3&3\end{pmatrix}\to\begin{pmatrix}1&1&-2\\0&1&-1\\0&0&0\end{pmatrix},$$

得线性无关的一个特征向量为 $X_{21}=(1,1,1)^{\mathrm{T}}$.

因此由定理 6.1 知 A 可对角化,且有

$$P=\begin{pmatrix}-1&-1&1\\0&1&1\\1&0&1\end{pmatrix},\quad P^{-1}AP=\begin{pmatrix}1&0&0\\0&1&0\\0&0&4\end{pmatrix}$$

可以验算结果的正确性,但验算时用 $AP=PA$ 更方便,不用求逆了.

(2) 对矩阵 $A=\begin{pmatrix}2&2&-1\\-1&0&1\\-1&0&2\end{pmatrix}$,我们在前一节已经算过它的特征值与特征向量,特征值 $\lambda_1=1$(2 重根),$\lambda_2=2$.

对应 $\lambda_1=1$ 只得到一个线性无关的特征向量 $X_1=(1,0,1)^{\mathrm{T}}$. 而 $\lambda_2=2$,有一个线性无关的特征向量 $X_2=(0,1,2)^{\mathrm{T}}$. A 的所有特征向量要么是 X_1 的倍数,要么是 X_2 的倍数,故 A 的线性无关的特征向量最多是 2 个,由定理 6.1 知,A 不可对角化.

下面,我们总结一下解决一个矩阵能否对角化问题的步骤:

(1) 求 A 的特征值. 设 A 的特征多项式为

$$f_A(\lambda)=|\lambda I-A|=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s},$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 互不相同.

(2) 对每个特征值 λ_i 计算特征向量,即对矩阵 $\lambda_i I-A$ 作高斯消元法,$(\lambda_i I-A)\to$阶梯形.

求出 λ_i 的所有线性无关的特征向量,设为

$$X_{i1},\quad X_{i2},\quad \cdots,\quad X_{im_i}.$$

如果 $m_1+m_2+\cdots+m_s<n$,则 A 不可对角化,事情就不必进行下去了. 如果 $m_1+m_2+\cdots+m_s=n$,则 A 可对角化,进行下一步.

(3) 令 $P=(X_{11},\cdots,X_{1n_1},\cdots,X_{s1},\cdots,X_{sn_s})$,则有

$$P^{-1}AP=\mathrm{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{n_1\text{个}},\cdots,\underbrace{\lambda_s,\cdots,\lambda_s}_{n_s\text{个}}).$$

注 首先,说明一下 $m_1+m_2+\cdots+m_s$ 是线性无关的特征向量的最大个数. 如上讨论,X_{i1},

$X_{i2},\cdots,X_{im_i}$ 是方程组 $(\lambda_i I-A)X=0$ 的基础解系.又由不同特征值下的特征向量是线性无关的,可知向量组

$$X_{11},\cdots,X_{1m_1},X_{21},\cdots,X_{2m_2},\cdots,X_{s1},\cdots,X_{sm_s} \qquad (*)$$

是线性无关的.对 A 的任意特征向量 X,设 X 属于特征值 λ_{i_0},于是 X 可由 $X_{i_01},\cdots,X_{i_0m_{i_0}}$ 线性表出,从而,向量组(*)是全体特征向量的一个极大无关组,$m_1+m_2+\cdots+m_s$ 是线性无关的特征向量的最大个数.简而言之,对每个特征值 λ_i,找出个数最多的属于 λ_i 的线性无关的特征向量组(局部上),并起来就得到个数最多的线性无关的特征向量组(整体上).

其次,对于每个特征值 λ_i,m_i 最大可取到多大呢?可以证明,m_i 的上界是 n_i,即 $m_i\leqslant n_i$ $(\forall i)$.其中 n_i 是 λ_i 作为特征多项式 $f_A(\lambda)$ 的根的重数.(出于篇幅与结构的考虑,本书略去证明.)又因 $n_1+n_2+\cdots+n_s=n$,所以,A 可对角化当且仅当 $m_i=n_i(\forall i)$.

6.2.3 矩阵对角化的应用与例

矩阵对角化的一个重要应用是用来简化矩阵的乘方的计算:如果 A 可对角化,即存在可逆阵 P 使

$$P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n),$$

两边作 k 次方,因而

$$(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)=P^{-1}A^kP=\mathrm{diag}(\lambda_1^k,\lambda_2^k,\cdots,\lambda_n^k),$$

得到计算矩阵的乘方的公式:

$$A^k=P\mathrm{diag}(\lambda_1^k,\lambda_2^k,\cdots,\lambda_n^k)P^{-1}.$$

例 6.4 设 $A=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}$,计算 A^n.

解 利用对角化来简化计算.由例 6.3 中已得到

$$P^{-1}AP=\begin{pmatrix}1&0&0\\0&1&0\\0&0&4\end{pmatrix},$$

两边 n 次方,得到

$$P^{-1}A^nP=\begin{pmatrix}1&0&0\\0&1&0\\0&0&4^n\end{pmatrix},$$

故

$$A^n=P\begin{pmatrix}1&0&0\\0&1&0\\0&0&4^n\end{pmatrix}P^{-1},$$

把 $\boldsymbol{P}=\begin{pmatrix}-1 & -1 & 1\\ 0 & 1 & 1\\ 1 & 0 & 1\end{pmatrix}$代入,可求得

$$\boldsymbol{A}^n=\boldsymbol{P}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 4^n\end{pmatrix}\boldsymbol{P}^{-1}=\frac{1}{3}\begin{pmatrix}2+4^n & -1+4^n & -1+4^n\\ -1+4^n & 2+4^n & -1+4^n\\ -1+4^n & -1+4^n & 2+4^n\end{pmatrix}.$$

例 6.5 若 $\boldsymbol{A}$ 是 n 阶可逆阵,可对角化,则 $\boldsymbol{A}^{-1},\boldsymbol{A}^*,\boldsymbol{A}^k(k\in\mathbb{Z}^+)$和 $f(\boldsymbol{A})$(f 为实多项式)也可对角化.

证明 (1)由于 $\boldsymbol{A}$ 可对角化,则有可逆阵 $\boldsymbol{P}$ 使

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n). \tag{6.14}$$

又由于 $\boldsymbol{A}$ 可逆,$\lambda_1,\lambda_2,\cdots,\lambda_n$ 均不为 0.对(6.14)两边求逆,得

$$\boldsymbol{P}^{-1}\boldsymbol{A}^{-1}\boldsymbol{P}=\operatorname{diag}(\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}).$$

所以 $\boldsymbol{A}^{-1}$可对角化.

(2)由 $\boldsymbol{A}\boldsymbol{A}^*=|\boldsymbol{A}|\boldsymbol{I}$ 及 $\boldsymbol{A}$ 可逆,得 $\boldsymbol{A}^*=|\boldsymbol{A}|\boldsymbol{A}^{-1}$和

$$\boldsymbol{P}^{-1}\boldsymbol{A}^*\boldsymbol{P}=|\boldsymbol{A}|\operatorname{diag}(\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}).$$

所以 $\boldsymbol{A}^*$ 可对角化.

(3)对式(6.14)两边作整数幂的运算,得

$$\boldsymbol{P}^{-1}\boldsymbol{A}^k\boldsymbol{P}=\operatorname{diag}(\lambda_1^k,\lambda_2^k,\cdots,\lambda_n^k)\quad(k\in\mathbb{Z}^+).$$

(4)$\boldsymbol{P}^{-1}f(\boldsymbol{A})\boldsymbol{P}=\operatorname{diag}(f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n))$. □

6.2.4 特征值理论的几个应用

下面,来看三个具体的实例.通过它们,读者可了解特征值理论的实用意义.

例 6.6 人员流动问题.

(案例问题)某试验性生产线每年初进行熟练工与非熟练工的人数统计,然后将$\frac{1}{6}$熟练工支援其他生产部门(成为其他部门的非熟练工),其缺额由招收新的非熟练工补齐.非熟练工经过培训及实践至年终考核有$\frac{2}{5}$成为熟练工.假设第一年初统计的熟练工和非熟练工各占一半,求以后每年初统计的熟练工和非熟练工所占的百分比.

(模型建立)设第 n 年初的熟练工和非熟练工所占的百分比分别为 x_n 和 y_n,记为向量$\begin{pmatrix}x_n\\ y_n\end{pmatrix}$.因为第一年统计的熟练工和非熟练工各占一半,所以$\begin{pmatrix}x_1\\ y_1\end{pmatrix}=\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix}$.为了求出以后每年 1

月份统计的熟练工和非熟练工所占的百分比，即求$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix}$与$\begin{pmatrix} x_n \\ y_n \end{pmatrix}$的关系式，然后再根据这个关系式求$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix}$.

根据已知条件可得

$$x_{n+1} = \left(1 - \frac{1}{6}\right) x_n + \frac{2}{5}\left(\frac{1}{6}x_n + y_n\right) = \frac{9}{10}x_n + \frac{2}{5}y_n,$$

$$y_{n+1} = \left(1 - \frac{2}{5}\right)\left(\frac{1}{6}x_n + y_n\right) = \frac{1}{10}x_n + \frac{3}{5}y_n,$$

即
$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} \frac{9}{10} & \frac{2}{5} \\ \frac{1}{10} & \frac{3}{5} \end{pmatrix}\begin{pmatrix} x_n \\ y_n \end{pmatrix}.$$

令 $\boldsymbol{A} = \begin{pmatrix} \frac{9}{10} & \frac{2}{5} \\ \frac{1}{10} & \frac{3}{5} \end{pmatrix}$，则

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \boldsymbol{A}\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \cdots = \boldsymbol{A}^n\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}.$$

（模型求解）方法一：计算 $\boldsymbol{A}^n$. 先将 $\boldsymbol{A}$ 对角化，这样可以简化计算

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} \frac{9}{10} - \lambda & \frac{2}{5} \\ \frac{1}{10} & \frac{3}{5} - \lambda \end{vmatrix} = (\lambda - 1)\left(\lambda - \frac{1}{2}\right).$$

由此得矩阵 $\boldsymbol{A}$ 的两个特征值 $\lambda_1 = 1, \lambda_2 = \frac{1}{2}$.

当 $\lambda_1 = 1$ 时，解 $(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$ 得对应于 $\lambda_1 = 1$ 的一个特征向量 $\boldsymbol{\xi}_1 = (4,1)^{\mathrm{T}}$.

当 $\lambda_2 = \frac{1}{2}$ 时，解 $\left(\boldsymbol{A} - \frac{1}{2}\boldsymbol{I}\right)\boldsymbol{X} = \boldsymbol{0}$ 得对应于 $\lambda_2 = \frac{1}{2}$ 的一个特征向量 $\boldsymbol{\xi}_2 = (-1,1)^{\mathrm{T}}$.

令 $\boldsymbol{P} = \begin{pmatrix} 4 & -1 \\ 1 & 1 \end{pmatrix}$，则

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

所以 $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}, \boldsymbol{A}^n = \boldsymbol{P}\boldsymbol{\Lambda}^n\boldsymbol{P}^{-1}$，则

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \boldsymbol{A}^n \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \boldsymbol{P}\boldsymbol{\Lambda}^n\boldsymbol{P}^{-1}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

$$= \begin{pmatrix} 4 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \left(\frac{1}{2}\right)^n \end{pmatrix} \begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ -\frac{1}{5} & \frac{4}{5} \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} \dfrac{4-3\times 2^{-n-1}}{5} \\ \dfrac{1+3\times 2^{-n-1}}{5} \end{pmatrix}.$$

(模型分析)当 $n\to\infty$ 时，$\dfrac{4-3\times2^{-n-1}}{5}\to\dfrac{4}{5}$，$\dfrac{1+3\times2^{-n-1}}{5}\to\dfrac{1}{5}$.

这说明，随着 n 的增加，熟练工和非熟练工所占的百分比趋于稳定，分别趋向于 80% 和 20%.

方法二：将初始向量 $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ 表示为两个特征向量 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ 的线性组合，设组合系数为 k_1, k_2，即

$$k_1\boldsymbol{\xi}_1 + k_2\boldsymbol{\xi}_2 = \begin{pmatrix} 4 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix},$$

求解上述关于 k_1, k_2 的线性方程组得

$$\begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{5} \\ \frac{3}{10} \end{pmatrix}.$$

故有

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \boldsymbol{A}^n \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \boldsymbol{A}^n(k_1\boldsymbol{\xi}_1 + k_2\boldsymbol{\xi}_2) = k_1\lambda_1^n\boldsymbol{\xi}_1 + k_2\lambda_2^n\boldsymbol{\xi}_2$$

$$= \frac{1}{5}\begin{pmatrix} 1 \\ 4 \end{pmatrix} + \frac{3}{10\cdot 2^n}\begin{pmatrix} -1 \\ 1 \end{pmatrix} \to \begin{pmatrix} \frac{1}{5} \\ \frac{4}{5} \end{pmatrix} \quad (n\to\infty \text{ 时}).$$

例 6.7 色盲基因的发展趋势.

(模型准备)伴性基因是一种位于 X 染色体上的基因.例如，红绿色盲基因是一种隐性的伴性基因.令 $x_1^{(0)}>0$ 为男性中有色盲基因的比例,并令 $x_2^{(0)}>0$ 为女性中有色盲基因的比例.由于男性从母亲获得一个 X 染色体,并且不从父亲处获得 X 染色体,所以下一代的男性中色盲的比例 $x_1^{(1)}$ 将和上一代的女性中所含有隐性色盲基因比例相同.由于女性从双亲处分别得到一个 X 染色体,所以下一代女性中含有隐性基因的比例 $x_2^{(1)}$ 将为 $x_1^{(0)}$ 和 $x_2^{(0)}$ 的平均值.验证当代数增加时,男性和女性中含有色盲基因的比例将趋于相同的数值.

(模型建立)设 $x_1^{(k)}$ 为第 k+1 代男性中有色盲基因的比例,$x_2^{(k)}$ 为第 k+1 代女性中有色盲基因的比例,则第一代男性和女性中有色盲基因的比例分别为 $x_1^{(0)}$ 和 $x_2^{(0)}$:

$$\begin{cases} x_1^{(k)} = x_2^{(k-1)}, \\ x_2^{(k)} = \dfrac{x_1^{(k-1)} + x_2^{(k-1)}}{2}. \end{cases} \tag{6.15}$$

于是问题就转化为把 $x_1^{(k)}, x_2^{(k)}$ 表示成 k 的函数,并确定 $\lim\limits_{k\to\infty} x_1^{(k)} = \lim\limits_{k\to\infty} x_2^{(k)}$.

(模型求解)方法一:由式(6.15),于是有

$$\begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix} \begin{pmatrix} x_1^{(k-1)} \\ x_2^{(k-1)} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix}^k \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \end{pmatrix}.$$

记 $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ \dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix}$, 要计算 $\boldsymbol{A}^k$,先将 $\boldsymbol{A}$ 对角化.

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} -\lambda & 1 \\ \dfrac{1}{2} & \dfrac{1}{2} - \lambda \end{vmatrix} = \left(\lambda - \frac{1}{2}\right)\lambda - \frac{1}{2} = (\lambda - 1)\left(\lambda + \frac{1}{2}\right),$$

由此可得 $\boldsymbol{A}$ 的两个特征值为 $\lambda_1 = -\dfrac{1}{2}, \lambda_2 = 1$.

解$\left(\boldsymbol{A}+\dfrac{1}{2}\boldsymbol{I}\right)\boldsymbol{X}=\boldsymbol{0}$ 得对应于 $\lambda_1 = -\dfrac{1}{2}$的一个特征向量 $\boldsymbol{\xi}_1 = (-2,1)^{\mathrm{T}}$.

解$(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 得对应于 $\lambda_2 = 1$ 的一个特征向量 $\boldsymbol{\xi}_2 = (1,1)^{\mathrm{T}}$.

令 $\boldsymbol{P} = \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix}$,则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda} = \begin{pmatrix} -\dfrac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}$.

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}, \quad \boldsymbol{A}^k = \boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1}.$$

因为

$$\boldsymbol{P}^{-1} = -\frac{1}{3}\begin{pmatrix} 1 & -1 \\ -1 & -2 \end{pmatrix},$$

所以
$$A^k=\begin{pmatrix}-2&1\\1&1\end{pmatrix}\begin{pmatrix}\left(-\frac{1}{2}\right)^k&0\\0&1\end{pmatrix}\begin{pmatrix}-\frac{1}{3}&\frac{1}{3}\\\frac{1}{3}&\frac{2}{3}\end{pmatrix}$$
$$=\begin{pmatrix}\frac{2}{3}\left(\frac{-1}{2}\right)^k+\frac{1}{3}&-\frac{2}{3}\left(\frac{-1}{2}\right)^k+\frac{2}{3}\\-\frac{1}{3}\left(\frac{-1}{2}\right)^k+\frac{1}{3}&\frac{1}{3}\left(\frac{-1}{2}\right)^k+\frac{2}{3}\end{pmatrix},$$

则
$$\begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\end{pmatrix}=A^k\begin{pmatrix}x_1^{(0)}\\x_2^{(0)}\end{pmatrix}=\begin{pmatrix}\frac{2}{3}\left(\frac{-1}{2}\right)^k+\frac{1}{3}&-\frac{2}{3}\left(-\frac{1}{2}\right)^k+\frac{2}{3}\\-\frac{1}{3}\left(\frac{-1}{2}\right)^k+\frac{1}{3}&\frac{1}{3}\left(\frac{-1}{2}\right)^k+\frac{2}{3}\end{pmatrix}\begin{pmatrix}x_1^{(0)}\\x_2^{(0)}\end{pmatrix},$$
$$\lim_{k\to\infty}\begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\end{pmatrix}=\begin{pmatrix}\frac{1}{3}&\frac{2}{3}\\\frac{1}{3}&\frac{2}{3}\end{pmatrix}\begin{pmatrix}x_1^{(0)}\\x_2^{(0)}\end{pmatrix}=\begin{pmatrix}\frac{1}{3}x_1^{(0)}+\frac{2}{3}x_2^{(0)}\\\frac{1}{3}x_1^{(0)}+\frac{2}{3}x_2^{(0)}\end{pmatrix}.$$

方法二:将初始向量$\begin{pmatrix}x_1^{(0)}\\x_2^{(0)}\end{pmatrix}$表示为两个特征向量的线性组合,即
$$\begin{pmatrix}x_1^{(0)}\\x_2^{(0)}\end{pmatrix}=k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2\quad(\exists k_1,k_2\in\mathbb{R}).$$
由 $x_1^{(0)},x_2^{(0)}>0$ 知,$k_2\neq0$. 于是有
$$\begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\end{pmatrix}=A^k\begin{pmatrix}x_1^{(0)}\\x_2^{(0)}\end{pmatrix}=A^k(k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2)=k_1\lambda_1^k\boldsymbol{\xi}_1+k_2\lambda_2^k\boldsymbol{\xi}_2$$
$$=\left(-\frac{1}{2}\right)^k k_1\begin{pmatrix}-2\\1\end{pmatrix}+k_2\begin{pmatrix}1\\1\end{pmatrix}\to\begin{pmatrix}k_2\\k_2\end{pmatrix}\quad(\text{当 }k\to\infty\text{ 时}).$$

因此,不论初始情形如何,从结果来看,当代数增加时,男性和女性中含有色盲基因的比例将趋于相同的数值.

注 由例 6.6、例 6.7,我们发现:对此类问题,随着 n 的增大,状态向量
$$\boldsymbol{x}_n=\boldsymbol{A}^n\boldsymbol{x}_0$$
会逐步趋近于某一个特征向量.该特征向量一定是属于绝对值相对较大的一个特征值.因此,绝对值最大的特征值及其特征向量,反映了此类问题的长期发展结果.例如,由例 6.6 我们得图 6.1.

此外,上述两例中我们均用了两种方法,但方法二相对比方法一简单,所以下面的例子

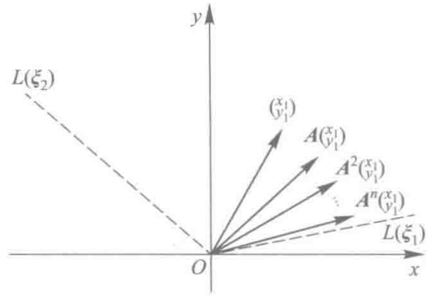


图 6.1

中，我们只用方法二来求解.

例 6.8　简单的种群增长问题.

（案例问题）经过统计，某地区猫头鹰和森林鼠的数量具有如下规律：每个月只有一半的猫头鹰可以存活；老鼠的数量每个月会增加 10%. 如果老鼠充足（数量为 R），则下个月猫头鹰的数量将会增加 $0.4R$. 平均每个月每只猫头鹰的捕食会导致 104 只老鼠死亡. 试确定该系统的演化情况.

（模型准备）不考虑其他因素对猫头鹰和森林鼠的数量的影响.

（模型建立）设猫头鹰和森林鼠在时刻 k 的数量为 $\boldsymbol{x}_k=\begin{pmatrix}O_k\\R_k\end{pmatrix}$，其中 k 是以月份为单位的时间，O_k 为研究区域中猫头鹰的数量（只），R_k 为老鼠的数量（千只），则

$$O_{k+1}=0.5O_k+0.4R_k,$$
$$R_{k+1}=-0.104O_k+1.1R_k.$$

分析 $\boldsymbol{x}_k$ 的变化趋势.

（模型求解）我们直接用方法二，令 $\boldsymbol{A}=\begin{pmatrix}0.5 & 0.4\\-0.104 & 1.1\end{pmatrix}$，则矩阵 $\boldsymbol{A}$ 的特征值 $\lambda_1=1.02$，$\lambda_2=0.58$，对应的特征向量 $\boldsymbol{v}_1=\begin{pmatrix}10\\13\end{pmatrix}$，$\boldsymbol{v}_2=\begin{pmatrix}5\\1\end{pmatrix}$.

初始向量 $\boldsymbol{x}_0$ 可以写成 $\boldsymbol{x}_0=c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2$，于是对于 $k\geqslant 0$，

$$\begin{aligned}\boldsymbol{x}_k&=c_1 1.02^k\boldsymbol{v}_1+c_2 0.58^k\boldsymbol{v}_2\\&=c_1 1.02^k\begin{pmatrix}10\\13\end{pmatrix}+c_2 0.58^k\begin{pmatrix}5\\1\end{pmatrix}.\end{aligned}$$

当 $k\to\infty$ 时，$0.58^k\to 0$. 假定 $c_1>0$，则对于所有足够大的 k，$\boldsymbol{x}_k$ 近似地等于 $c_1 1.02^k\boldsymbol{v}_1$，写成

$$\boldsymbol{x}_k\approx c_1 1.02^k\begin{pmatrix}10\\13\end{pmatrix}.$$

k 越大,近似程度越高,所以对于充分大的 k,

$$x_{k+1} \approx c_1 1.02^{k+1}\begin{pmatrix}10\\13\end{pmatrix} = 1.02x_k. \tag{6.16}$$

(模型分析)式(6.16)表明,最后猫头鹰和老鼠的数量几乎每个月都近似增加到原来的 1.02 倍,即有 2%的月增长率.O_k 与 R_k 的比值约为 10 : 13,即每 10 只猫头鹰对应着约 13 000 只老鼠.

§6.3 实对称阵的对角化

实对称阵是一类非常重要的矩阵,它有很好的性质:如它的特征值均为实数,且必然可对角化.另一方面,实对称阵有很多应用,下一节我们将利用实对称阵,讨论三维空间$\mathbb{R}^3$中的二次曲面分类问题.

首先,在数学上定义实对称矩阵.设 $\boldsymbol{A} \in M_n(\mathbb{R})$,满足 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为实对称阵.由 $\mathbb{R} \subseteq \mathbb{C}$,实对称矩阵也可以在复数域中给出定义.为此,先介绍复共轭的概念.

若 $z=a+bi(a,b \in \mathbb{R}, i=\sqrt{-1})$,定义 z 的共轭复数为 $\bar{z} \triangleq a-bi$.并且对任意复矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}, a_{ij} \in \mathbb{C}$,定义 $\boldsymbol{A}$ 的复共轭矩阵为 $\overline{\boldsymbol{A}}=(\bar{a}_{ij})_{m\times n}$.这时,如果 $\boldsymbol{A} \in M_n(\mathbb{C})$,使得 $\overline{\boldsymbol{A}}=\boldsymbol{A}$ 且 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为实对称矩阵.

容易验证,复共轭运算可与复数的加、减、乘、除四则运算交换顺序,即

$$\bar{z} \pm \bar{w} = \overline{z \pm w} \quad \bar{z}\cdot\bar{w} = \overline{z\cdot w} \quad \frac{\bar{z}}{\bar{w}} = \overline{\left(\frac{z}{w}\right)}.$$

另有如下正定性质:

$$z\cdot\bar{z} = a^2+b^2 \geqslant 0, \text{且等号成立当且仅当 } z=0.$$

6.3.1 实对称阵的特征值与特征向量的性质

首先实对称矩阵的特征值有如下定理.

定理 6.2 实对称矩阵的特征值是实数.

证明 设 $\boldsymbol{A}$ 是 n 阶实对称方阵,λ 是 $\boldsymbol{A}$ 的任意一个特征值,$\boldsymbol{X}$ 是 λ 的一个特征向量,即

$$\boldsymbol{AX} = \lambda\boldsymbol{X}, \tag{6.17}$$

将式(6.17)两边左乘 $\overline{\boldsymbol{X}}^{\mathrm{T}}$,得

$$\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{AX} = \lambda\overline{\boldsymbol{X}}^{\mathrm{T}}\boldsymbol{X}, \tag{6.18}$$

其中 $\overline{\boldsymbol{X}}$ 表示将 $\boldsymbol{X}$ 的每一个分量取其共轭复数,即如果 $\boldsymbol{X}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$,则 $\overline{\boldsymbol{X}}=(\bar{x}_1,\bar{x}_2,\cdots,\bar{x}_n)^{\mathrm{T}}$.

对式(6.18)两边求转置共轭,左端 $\overline{X}^{\mathrm{T}}AX$ 作转置共轭运算得 $\overline{(\overline{X}^{\mathrm{T}}AX)}^{\mathrm{T}}=\overline{X}^{\mathrm{T}}AX$(由于 A 是 n 阶实对称方阵,有 $\overline{A}^{\mathrm{T}}=A$);右端作转置共轭运算得 $\overline{(\lambda\overline{X}^{\mathrm{T}}X)}^{\mathrm{T}}=\overline{\lambda}\overline{X}^{\mathrm{T}}X$,因而有

$$\overline{X}^{\mathrm{T}}AX=\overline{\lambda}\overline{X}^{\mathrm{T}}X, \tag{6.19}$$

把式(6.18)和式(6.19)相减得

$$(\overline{\lambda}-\lambda)\overline{X}^{\mathrm{T}}X=0.$$

由 $\overline{X}^{\mathrm{T}}X>0$,故有 $\overline{\lambda}-\lambda=0$,即 $\overline{\lambda}=\lambda$,所以 λ 是实数. □

其实我们可直接从式(6.18)看出 λ 是实数:由于 $\overline{X}^{\mathrm{T}}X>0$ 是一个正实数,式(6.18)的左端 $\overline{X}^{\mathrm{T}}AX$ 是一个数,可看做是一个一阶矩阵,对它作转置共轭运算:$\overline{(\overline{X}^{\mathrm{T}}AX)}^{\mathrm{T}}=\overline{X}^{\mathrm{T}}AX$,故 $\overline{X}^{\mathrm{T}}AX$ 也是一个实数,所以

$$\lambda=\frac{\overline{X}^{\mathrm{T}}AX}{\overline{X}^{\mathrm{T}}X}$$

是实数.

至于实对称矩阵的特征向量,由于在求特征向量时所解的齐次方程组是实系数方程组,所以可以得到实的特征向量,而且对应不同特征值的特征向量还有如下正交关系.

定理 6.3 设 A 是 n 阶实对称方阵,λ_1,λ_2 是 A 的两个不同的特征值,X_1,X_2 是分别属于 λ_1,λ_2 的特征向量,则 $X_1\perp X_2$.

证明 由已知条件有

$$AX_1=\lambda_1X_1, \tag{6.20}$$

$$AX_2=\lambda_2X_2, \tag{6.21}$$

将式(6.21)两边左乘 X_1^{T},得

$$X_1^{\mathrm{T}}AX_2=\lambda_2X_1^{\mathrm{T}}X_2, \tag{6.22}$$

将式(6.20)两边转置,注意到 $A^{\mathrm{T}}=A$,得

$$X_1^{\mathrm{T}}A=\lambda_1X_1^{\mathrm{T}}, \tag{6.23}$$

将式(6.23)两边右乘 X_2,得

$$X_1^{\mathrm{T}}AX_2=\lambda_1X_1^{\mathrm{T}}X_2, \tag{6.24}$$

将式(6.22)减式(6.24),得

$$(\lambda_2-\lambda_1)X_1^{\mathrm{T}}X_2=0,$$

由于 $\lambda_1\neq\lambda_2$,即 $\lambda_1-\lambda_2\neq0$,所以

$$X_1^{\mathrm{T}}X_2=0,$$

即 $X_1\perp X_2$. □

例如,在例 6.3 中,$A=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}$,求得 $\lambda_1=1$(重数为 2),$\lambda_2=4$.对特征值 $\lambda_1=1$ 线性无

关的两个特征向量为 $X_{11}=(-1,0,1)^T, X_{12}=(-1,1,0)^T$.

对特征值 $\lambda_2=4$ 线性无关的一个特征向量为 $X_{21}=(1,1,1)^T$. 由于 $(X_{21},X_{11})=(X_{21},X_{12})=0$, 所以 $X_{21}\perp X_{11}, X_{21}\perp X_{12}$.

6.3.2 实对称阵的对角化

本节的主要结果为:

定理 6.4 设 A 是 n 阶实对称方阵, 则有正交阵 Q 使 $Q^{-1}AQ=\Lambda$, 其中 $\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$, $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 A 的全部特征值; $Q=(X_1,\cdots,X_n)$, $X_1,X_2,\cdots,X_n$ 为对应的特征向量.

证明 对 n 作归纳法.

$$n=1, A=(a_{11}), \text{可认为已对角化}, a_{11}\text{ 是特征值};$$

$$Q=I=(1), \text{是一阶正交阵}, X=(1)\text{ 是特征向量}.$$

假设结论对 $n-1$ 成立, 要证对 n 也成立.

由于 A 的所有特征值均为实数, 并有实的特征向量, 所以可以在欧氏空间 $\mathbb{R}^n$ 中讨论问题. 设 λ_1 是 A 的一个特征值, X_1 是对应的一个单位特征向量. 把 X_1 扩大为一组标准正交基: $X_1,X_2,\cdots,X_n$.

令 $Q_1=(X_1,X_2,\cdots,X_n)$, 作运算

$$\begin{aligned} AQ_1 &= (AX_1,AX_2,\cdots,AX_n)=(\lambda_1X_1,AX_2,\cdots,AX_n) \\ &= (X_1,X_2,\cdots,X_n)\begin{pmatrix}\lambda_1 & \boldsymbol{\alpha} \\ \mathbf{0} & A_1\end{pmatrix}=Q_1\begin{pmatrix}\lambda_1 & \boldsymbol{\alpha} \\ \mathbf{0} & A_1\end{pmatrix}. \end{aligned}$$

其中 $\boldsymbol{\alpha}$ 是一个 $n-1$ 维的行向量, A_1 是一个 $n-1$ 阶方阵. 因而得到

$$Q_1^{-1}AQ_1=\begin{pmatrix}\lambda_1 & \boldsymbol{\alpha} \\ \mathbf{0} & A_1\end{pmatrix}, \tag{6.25}$$

注意到 Q_1 是正交阵, $Q_1^{-1}=Q_1^T$, 式(6.25)的左边 $Q_1^{-1}AQ_1=Q_1^TAQ_1$ 是对称阵, 因而得到 $\boldsymbol{\alpha}=\mathbf{0}$, A_1 是一个 $n-1$ 阶实对称阵. 于是由归纳假设, 存在 $n-1$ 阶正交阵 Q_2, 使 $Q_2^{-1}A_2Q_2=\Lambda_2=\mathrm{diag}(\lambda_2,\cdots,\lambda_n)$, 其中 $\lambda_2,\cdots,\lambda_n$ 是 A_1 的全部特征值, 也是 A 的特征值.

令 $Q=Q_1\begin{pmatrix}1 & \mathbf{0} \\ \mathbf{0} & Q_2\end{pmatrix}$, $\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$, 显然 Q 是正交阵, 则有

$$\begin{aligned} Q^{-1}AQ &= \begin{pmatrix}1 & \mathbf{0} \\ \mathbf{0} & Q_2^{-1}\end{pmatrix}Q_1^{-1}AQ_1\begin{pmatrix}1 & \mathbf{0} \\ \mathbf{0} & Q_2\end{pmatrix} \\ &= \begin{pmatrix}1 & \mathbf{0} \\ \mathbf{0} & Q_2^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & \boldsymbol{\alpha} \\ \mathbf{0} & A_1\end{pmatrix}\begin{pmatrix}1 & \mathbf{0} \\ \mathbf{0} & Q_2\end{pmatrix} \end{aligned}$$

$$=\begin{pmatrix}\lambda_1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{Q}_2^{-1}\boldsymbol{A}_1\boldsymbol{Q}_2\end{pmatrix}=\begin{pmatrix}\lambda_1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Lambda}_2\end{pmatrix}=\boldsymbol{\Lambda},$$

所以定理成立. □

定理 6.4 说明了实对称矩阵有 n 个线性无关的特征向量.因此,如果不要求用正交矩阵作相似变换,则可直接用由特征向量组成的矩阵来作相似变换.如果要求用正交矩阵作相似变换,则在求出所有特征向量后用施密特正交化方法进行正交化.下面举例来说明这一点.

6.3.3 实对称阵的对角化(主轴化)方法

在具体举例之前,我们把上一讲中介绍的矩阵对角化方法做一些修改,应用到实对称矩阵的对角化问题.

实对称矩阵对角化的计算步骤:

(1) 求 $\boldsymbol{A}$ 的特征值,设 $\boldsymbol{A}$ 的特征多项式为

$$f_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 互不相同.

(2) 对每个特征值 λ_i 计算特征向量.对矩阵 $\lambda_i\boldsymbol{I}-\boldsymbol{A}$ 作高斯消元法:

$$(\lambda_i\boldsymbol{I}-\boldsymbol{A})\rightarrow\text{阶梯形}.$$

求出 λ_i 的所有线性无关的特征向量,设为

$$\boldsymbol{X}_{i1},\quad \boldsymbol{X}_{i2},\quad \cdots,\quad \boldsymbol{X}_{in_i}.$$

由 6.2.2 节最后的注,可知 λ_i 必有 n_i 个线性无关的特征向量($i=1,2,\cdots,s$).

(3) 求变换矩阵.

如果不要求用正交阵作相似变换,则令 $\boldsymbol{P}=(\boldsymbol{X}_{11},\cdots,\boldsymbol{X}_{1n_1},\cdots,\boldsymbol{X}_{s1},\cdots,\boldsymbol{X}_{sn_s})$,则有

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\mathrm{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{n_1\text{个}},\cdots,\underbrace{\lambda_s,\cdots,\lambda_s}_{n_s\text{个}}).$$

如果要求用正交阵作相似变换,则对每个特征值 λ_i 的特征向量 $\boldsymbol{X}_{i1},\boldsymbol{X}_{i2},\cdots,\boldsymbol{X}_{in_i}$ 用施密特正交化方法进行正交化,再单位化,得到正交矩阵 $\boldsymbol{Q}$,并有

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{AQ}=\mathrm{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{n_1\text{个}},\cdots,\underbrace{\lambda_s,\cdots,\lambda_s}_{n_s\text{个}}).$$

这称为主轴化方法.

例 6.9 设 $\boldsymbol{A}=\begin{pmatrix}1&1&1&-1\\1&1&-1&1\\1&-1&1&1\\-1&1&1&1\end{pmatrix}$,用两种方法将其对角化.

解 (1) 求特征值.

$$f_A(\lambda)=|\lambda I-A|=\begin{vmatrix}\lambda-1 & -1 & -1 & 1\\ -1 & \lambda-1 & 1 & -1\\ -1 & 1 & \lambda-1 & -1\\ 1 & -1 & -1 & \lambda-1\end{vmatrix}$$

$$=\begin{vmatrix}0 & \lambda-2 & \lambda-2 & -\lambda^2+2\lambda\\ 0 & \lambda-2 & 0 & \lambda-2\\ 0 & 0 & \lambda-2 & \lambda-2\\ 1 & -1 & -1 & \lambda-1\end{vmatrix}=(\lambda-2)^3(\lambda+2),$$

所以得 $\lambda_1=2$(3 重根),$\lambda_2=-2$.

(2) 求特征向量.

对 $\lambda_1=2$,对 $2I-A$ 进行高斯消元法:

$$2I-A=\begin{pmatrix}1 & -1 & -1 & 1\\ -1 & 1 & 1 & -1\\ -1 & 1 & 1 & -1\\ 1 & -1 & -1 & 1\end{pmatrix}\rightarrow\begin{pmatrix}1 & -1 & -1 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix},$$

所以得线性无关的特征向量:

$$\boldsymbol{\alpha}_1=(1,0,0,-1)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(1,0,1,0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_3=(1,1,0,0)^{\mathrm{T}}.$$

对 $\lambda_2=-2$,对 $-2I-A$ 进行高斯消元法:

$$-2I-A=\begin{pmatrix}-3 & -1 & -1 & 1\\ -1 & -3 & 1 & -1\\ -1 & 1 & -3 & -1\\ 1 & -1 & -1 & -3\end{pmatrix}\rightarrow\begin{pmatrix}1 & -1 & -1 & -3\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0\end{pmatrix},$$

所以得线性无关的特征向量:

$$\boldsymbol{\alpha}_4=(1,-1,-1,1)^{\mathrm{T}}.$$

(3) 求变换矩阵.

令 $P=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)=\begin{pmatrix}1 & 1 & 1 & 1\\ 0 & 0 & 1 & -1\\ 0 & 1 & 0 & -1\\ -1 & 0 & 0 & 1\end{pmatrix}$,

则

$$P^{-1}AP=\begin{pmatrix}2 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & -2\end{pmatrix}.$$

下面求正交阵作为变换矩阵.

将 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 用施密特正交化方法正交化：

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1=(1,0,0,-1)^{\mathrm{T}},$$

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1=\left(\frac{1}{2},0,1,\frac{1}{2}\right)^{\mathrm{T}},$$

$$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2=\left(\frac{1}{3},1,-\frac{1}{3},\frac{1}{3}\right)^{\mathrm{T}}.$$

单位化得

$$\boldsymbol{\varepsilon}_1=\left(\frac{1}{\sqrt{2}},0,0,-\frac{1}{\sqrt{2}}\right)^{\mathrm{T}},$$

$$\boldsymbol{\varepsilon}_2=\left(\frac{1}{\sqrt{6}},0,\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}}\right)^{\mathrm{T}},$$

$$\boldsymbol{\varepsilon}_3=\left(\frac{1}{2\sqrt{3}},\frac{3}{2\sqrt{3}},\frac{-1}{2\sqrt{3}},\frac{1}{2\sqrt{3}}\right)^{\mathrm{T}},$$

$$\boldsymbol{\varepsilon}_4=\left(\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right)^{\mathrm{T}},$$

令 $\boldsymbol{Q}=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)=\begin{pmatrix}\frac{1}{\sqrt{2}}, & \frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & \frac{1}{2}\\ 0 & 0 & \frac{3}{2\sqrt{3}} & -\frac{1}{2}\\ 0 & \frac{2}{\sqrt{6}} & \frac{-1}{2\sqrt{3}} & -\frac{1}{2}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & \frac{1}{2}\end{pmatrix}$,

则 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\mathrm{diag}(2,2,2,-2)$.

至此，实对称矩阵的对角化问题圆满解决了.

思考题：设 $\boldsymbol{A}$ 是 n 阶实对称阵，则 $\boldsymbol{A}^*,\boldsymbol{A}^k(k\in\mathbb{Z}^+)$ 和 $f(\boldsymbol{A})$（f 为实多项式）可对角化.

提示：只需证明 $\boldsymbol{A}^*,\boldsymbol{A}^k(k\in\mathbb{Z}^+)$ 和 $f(\boldsymbol{A})$ 是实对称矩阵.

习　题　六

1. 求下列矩阵 $\boldsymbol{A}$ 的全部特征值和特征向量：

(1) $\boldsymbol{A}=\begin{pmatrix}2 & 2 & -2\\ 2 & 5 & -4\\ -2 & -4 & 5\end{pmatrix}$；　　(2) $\boldsymbol{A}=\begin{pmatrix}2 & 3 & 2\\ 1 & 8 & 2\\ -2 & -14 & -3\end{pmatrix}$；

（3）$A=\begin{pmatrix}2&-1&2\\5&-3&3\\-1&0&-2\end{pmatrix}$；　　（4）$A=\begin{pmatrix}0&\frac{1}{2}&\frac{1}{2}\\1&-\frac{1}{2}&\frac{1}{2}\\1&-\frac{1}{2}&\frac{1}{2}\end{pmatrix}$.

2. 设 A 为 2 阶矩阵，α_1,α_2 为线性无关的 2 维列向量，且 $A\alpha_1=0$，$A\alpha_2=2\alpha_1+\alpha_2$，求 A 的非零特征值.

3. 已知 $A=\begin{pmatrix}1&2\\-1&4\end{pmatrix}$，求 A^m（其中 m 是任一正整数）.

4. 设 3 阶实对称矩阵 A 的秩为 2，且 $\lambda_1=\lambda_2=6$ 是 A 的二重特征值. 若 $\alpha_1=(1,1,0)^{\mathrm{T}}$，$\alpha_2=(2,1,1)^{\mathrm{T}}$，$\alpha_3=(-1,2,-3)^{\mathrm{T}}$ 都是 A 的属于特征值 6 的特征向量.

（1）求 A 的另一特征值和对应的特征向量；

（2）求矩阵 A.

5. 已知矩阵 $A=\begin{pmatrix}-2&0&0\\2&x&2\\3&1&1\end{pmatrix}$ 与 $B=\begin{pmatrix}-1&0&0\\0&2&0\\0&0&y\end{pmatrix}$ 相似.

（1）求 x 与 y 的值；

（2）求可逆矩阵 P，使得 $P^{-1}AP=B$.

6. 设 A 为 3 阶矩阵，$\alpha_1,\alpha_2,\alpha_3$ 是线性无关的三维列向量，且满足 $A\alpha_1=\alpha_1+\alpha_2+\alpha_3$，$A\alpha_2=2\alpha_2+\alpha_3$，$A\alpha_3=2\alpha_2+3\alpha_3$.

（1）求矩阵 B，使得 $A(\alpha_1,\alpha_2,\alpha_3)=(\alpha_1,\alpha_2,\alpha_3)B$；

（2）求矩阵 A 的特征值.

7. 已知 3 阶矩阵 A 与三维向量 X,AX,A^2X 线性无关，且满足 $A^3X=3AX-2A^2X$.

（1）记 $P=(X,AX,A^2X)$，求 3 阶矩阵 B，使 $A=PBP^{-1}$；

（2）计算行列式 $|A+I|$.

8. 已知向量 $\xi=\begin{pmatrix}1\\1\\-1\end{pmatrix}$ 是矩阵 $A=\begin{pmatrix}2&-1&1\\5&a&3\\-1&b&-2\end{pmatrix}$ 的一个特征向量.

（1）确定参数 a,b 及特征向量 ξ 所对应的特征值 λ；

（2）请问 A 能否对角化？说明理由.

9. 设 3 阶实对称矩阵 A 的各行元素之和均为 3，向量 $\alpha_1=(-1,2,-1)^{\mathrm{T}}$，$\alpha_2=(0,-1,1)^{\mathrm{T}}$ 是线性方程组 $AX=0$ 的两个解.

（1）求 A 的特征值与特征向量；

(2) 求正交矩阵 $\boldsymbol{Q}$ 和对角矩阵 $\boldsymbol{\Lambda}$，使得 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{\Lambda}$；

(3) 求 $\boldsymbol{A}$ 及 $\left(\boldsymbol{A}-\dfrac{3}{2}\boldsymbol{I}\right)^{6}$，其中 $\boldsymbol{I}$ 为 3 阶单位矩阵.

第七章 二次型

§7.1 二次型和它的标准形

二次型是研究二次曲线和二次曲面的代数工具.此外,它在数学的许多分支以及物理学和工程技术中都有重要应用.首先,我们来学习任意数域 F 上的二次型理论.

定义 7.1 数域 F 上的 n 元**二次型**是系数在 F 中的 n 个不定元的**二次齐次多项式**.用 $x_1,x_2,\cdots,x_n$ 来表示它的 n 个不定元,则二次型的一般形式为

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2&+2a_{12}x_1x_2+2a_{13}x_1x_3+\cdots+2a_{1n}x_1x_n\\&+a_{22}x_2^2+2a_{23}x_2x_3+\cdots+2a_{2n}x_2x_n\\&+\cdots\\&+a_{nn}x_n^2,\end{aligned}\tag{7.1}$$

其中 $a_{ij}\in F,1\leqslant i\leqslant j\leqslant n$.

注意 表达式(7.1)中把交叉项 $x_ix_j(i<j)$ 的系数写成 $2a_{ij}$ 是为了后面讨论的方便.

当 $i>j$ 时令 $a_{ij}=a_{ji},1\leqslant j<i\leqslant n$,则式(7.1)也可以写成

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2&+a_{12}x_1x_2+a_{13}x_1x_3+\cdots+a_{1n}x_1x_n\\&+a_{21}x_2x_1+a_{22}x_2^2+a_{23}x_2x_3+\cdots+a_{2n}x_2x_n\\&+a_{31}x_3x_1+a_{32}x_3x_2+a_{33}x_3^2+\cdots+a_{3n}x_3x_n\\&+\cdots\\&+a_{n1}x_nx_1+a_{n2}x_nx_2+a_{n3}x_nx_3+\cdots+a_{nn}x_n^2\\=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,\end{aligned}\tag{7.2}$$

把(7.2)中的系数按原来的顺序排成一个 n 阶矩阵 $\boldsymbol{A}$,即

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}.$$

由于对任意 $1\leqslant i,j\leqslant n$ 有 $a_{ij}=a_{ji}$,所以矩阵 A 是对称矩阵,称 A 为二次型(7.1)的矩阵.显然,给定一个二次型 $f(x_1,x_2,\cdots,x_n)$,则它的矩阵 A 是唯一确定的.因为 A 的主对角线上的元素 a_{ii} 是 $f(x_1,x_2,\cdots,x_n)$ 中 x_i^2 的系数,而 $a_{ij}(i\neq j)$ 是 $f(x_1,x_2,\cdots,x_n)$ 中 $x_ix_j(i<j)$ 系数的一半.(注意到在 $f(x_1,x_2,\cdots,x_n)$ 的表达式中 x_ix_j 与 x_jx_i 是同一项,要合并在一起.)记

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

则二次型(7.1)也可以写成

$$f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX. \tag{7.3}$$

例 7.1 写出下列二次型的矩阵

(1) $f_1(x_1,x_2,x_3)=x_1^2-2x_1x_2+3x_1x_3-x_2x_3+4x_3^2$;

(2) $f_2(x_1,x_2,x_3,x_4)=x_1x_3+x_2x_4$.

解 由定义,这两个二次型的矩阵分别为

$$\begin{pmatrix} 1 & -1 & \frac{3}{2} \\ -1 & 0 & -\frac{1}{2} \\ \frac{3}{2} & -\frac{1}{2} & 4 \end{pmatrix} \quad 和 \quad \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \end{pmatrix}.$$

任给一 F 上的 n 元二次型.有一个 F 上的 n 阶对称矩阵(即此二次型的矩阵)跟它对应.反之,任给数域 F 上的一个 n 阶对称矩阵 A,也可以得到 F 上的一个 n 元二次型

$$f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX,$$

其中 $X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.这表明,在 F 上的 n 元二次型集合与 F 上的 n 阶对称矩阵集合之间有一个一

一对应关系.

令 $y_1,y_2,\cdots,y_n$ 为另一组不定元,记

$$\boldsymbol{Y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix},$$

设 $\boldsymbol{C}$ 是数域 $\boldsymbol{F}$ 上的一个 n 阶可逆矩阵,称关系式

$$\boldsymbol{X}=\boldsymbol{CY} \tag{7.4}$$

为不定元 $x_1,x_2,\cdots,x_n$ 到不定元 $y_1,y_2,\cdots,y_n$ 的一个非退化(可逆)线性替换.

定义 7.2 数域 F 上的两个二次型称为是等价的若存在一个非退化线性替换把一个变成另一个.

注 从多项式的观点来看,一个多项式的不定元使用什么字母来表示是无关紧要的,所以说

$$\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,\qquad \sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}y_iy_j$$

表示的是同一个二次型.

n 元二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$(其中 $\boldsymbol{A}$ 为对称矩阵)经过非退化线性替换 $\boldsymbol{X}=\boldsymbol{CY}$ 后变成

$$(\boldsymbol{CY})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{CY})=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{ACY},$$

记 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$,则上式为 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{BY}$,这是变量 $y_1,y_2,\cdots,y_n$ 的一个二次型,又由于

$$\boldsymbol{B}^{\mathrm{T}}=(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC})^{\mathrm{T}}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{C}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\boldsymbol{B},$$

即 $\boldsymbol{B}$ 也是对称矩阵.于是二次型 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{BY}$ 的矩阵正好是 $\boldsymbol{B}$.

定义 7.3 数域 F 上的两个 n 阶矩阵 $\boldsymbol{A},\boldsymbol{B}$ 称为是**合同的**,记为 $\boldsymbol{A}\simeq\boldsymbol{B}$,若存在 F 上的一个可逆矩阵 $\boldsymbol{C}$,使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\boldsymbol{B}$.

显然矩阵的合同满足反身性、对称性和传递性.

前文已经证明了若两个二次型等价,则它们的矩阵是合同的.反之,若二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 和 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{BY}$ 的矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 合同,则存在可逆矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\boldsymbol{B}$.从而非退化线性替换 $\boldsymbol{X}=\boldsymbol{CY}$ 就把二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 变成 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{BY}$,即这两个二次型等价.这样便证出下面结论.

命题 7.1 数域 F 上两个 n 元二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 和 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{BY}$ 等价当且仅当它们的矩阵,即 n 阶对称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 合同.

二次型理论中的一个基本问题是数域 F 上的二次型能不能等价于一个只含有平方项(即交叉项的系数全为 0)的二次型.显然,只含有平方项的二次型的矩阵是对角矩阵,反之亦然,因此这个问题用矩阵的语言来说就是对称矩阵能不能合同于一个对角矩阵.

首先证明本节的主要结论.

定理 7.1 数域 F 上任一对称矩阵 $\boldsymbol{A}$ 都合同于一个对角矩阵.

证明 对 $\boldsymbol{A}$ 的阶数 n 作归纳.

当 $n=1$ 时,$\boldsymbol{A}$ 本身就是对角矩阵.结论显然成立.假设任一 $n-1$ 阶对称矩阵都合同于对角矩阵,下面来看 n 阶对称矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$.

情形 1:$a_{11}\neq 0$. 将 $\boldsymbol{A}$ 写成分块矩阵的情形,

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & \boldsymbol{A}_1 \end{pmatrix}$$

其中 $\boldsymbol{\alpha}$ 为一 $1\times(n-1)$ 矩阵,$\boldsymbol{A}_1$ 为 $n-1$ 阶对称矩阵(因为 $\boldsymbol{A}$ 对称).

令 $\boldsymbol{C}_1=\begin{pmatrix} 1 & -a_{11}^{-1}\boldsymbol{\alpha} \\ \boldsymbol{0} & \boldsymbol{I}_{n-1} \end{pmatrix}$,则

$$\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}_1=\begin{pmatrix} 1 & \boldsymbol{0} \\ -a_{11}^{-1}\boldsymbol{\alpha}^{\mathrm{T}} & \boldsymbol{I}_{n-1} \end{pmatrix}\begin{pmatrix} a_{11} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & \boldsymbol{A}_1 \end{pmatrix}\begin{pmatrix} 1 & -a_{11}^{-1}\boldsymbol{\alpha} \\ \boldsymbol{0} & \boldsymbol{I}_{n-1} \end{pmatrix}$$

$$=\begin{pmatrix} a_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_2 \end{pmatrix},$$

其中 $\boldsymbol{A}_2=\boldsymbol{A}_1-a_{11}^{-1}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}$,由于

$$\boldsymbol{A}_2^{\mathrm{T}}=(\boldsymbol{A}_1-a_{11}^{-1}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha})^{\mathrm{T}}=\boldsymbol{A}_1^{\mathrm{T}}-a_{11}^{-1}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}=\boldsymbol{A}_1-a_{11}^{-1}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}=\boldsymbol{A}_2,$$

即 $\boldsymbol{A}_2$ 为 $n-1$ 阶对称矩阵,则由归纳假设,存在数域 F 上 $n-1$ 阶可逆矩阵 $\boldsymbol{C}_2$ 使得 $\boldsymbol{C}_2^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{C}_2$ 为对角矩阵.令

$$\boldsymbol{C}=\boldsymbol{C}_1\begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_2 \end{pmatrix},$$

则显然 $\boldsymbol{C}$ 为可逆矩阵,且

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_2^{\mathrm{T}} \end{pmatrix}\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}_1\begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_2 \end{pmatrix}$$

$$=\begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_2^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} a_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_2 \end{pmatrix}\begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_2 \end{pmatrix}$$

$$=\begin{pmatrix} a_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_2^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{C}_2 \end{pmatrix}$$

为对角矩阵,故 $\boldsymbol{A}$ 与对角矩阵合同.

情形 2:$a_{11}=0$,但存在 $2\leqslant i\leqslant n$,有 $a_{ii}\neq 0$. 令 $\boldsymbol{P}(1,i)^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}(1,i)=\boldsymbol{B}$,其中 $\boldsymbol{P}(1,i)$ 为交换单位矩阵第 1 列和第 i 列所得的初等矩阵.故 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同.又 $\boldsymbol{B}$ 的第 1 行第 1 列元素为 $a_{ii}\neq 0$. 根据情形 1,$\boldsymbol{B}$ 与对角矩阵合同,由合同的传递性得 $\boldsymbol{A}$ 与对角矩阵合同.

情形 3:$a_{11}=a_{22}=\cdots=a_{nn}=0$.

(i) 若 $\boldsymbol{A}=\boldsymbol{O}$.则 A 已为对角矩阵.

(ii) 若 $\boldsymbol{A}\neq\boldsymbol{O}$.则存在 $a_{ij}\neq 0, i\neq j$,显然 $\boldsymbol{P}(j,i(1))^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}(j,i(1))=\boldsymbol{H}$ 的第 i 个主对角元为 $2a_{ij}\neq 0$,其中 $\boldsymbol{P}(j,i(1))$ 为把单位矩阵第 j 列加到第 i 列上所得的初等矩阵,故 $\boldsymbol{A}$ 与 $\boldsymbol{H}$ 合同.又由情形 2 知 $\boldsymbol{H}$ 与对角矩阵合同.仍由合同的传递性知 $\boldsymbol{A}$ 与对角矩阵合同.

由归纳法原理,结论对任意正整数 n 均成立,即任一 n 阶对称矩阵都合同于一个对角矩阵. □

从定理 7.1 立得

定理 7.2 数域 F 上任一二次型都等价于一个只含平方项的二次型.

对于给定的对称矩阵 $\boldsymbol{A}$,如何求出与 $\boldsymbol{A}$ 合同的对角矩阵,又如何求出可逆矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 为对角矩阵?

由定理 7.1 的证明,可把 n 阶情形化为 $n-1$ 阶情形,再化为 $n-2$ 阶情形,一直进行下去.在每次降阶的过程中,若矩阵不为 $\boldsymbol{O}$,可以先把它的第 1 行第 1 列位置元素化为非 0. 再利用初等列变换把第 1 行中除第 1 列外的其他元素都化为 0. 这些初等列变换对应的即为可逆矩阵 $\boldsymbol{C}_1=\begin{pmatrix}1 & -a_{11}^{-1}\boldsymbol{\alpha}\\ \boldsymbol{0} & \boldsymbol{I}_{n-1}\end{pmatrix}$.在矩阵的左端乘 $\boldsymbol{C}_1^{\mathrm{T}}$,就把矩阵的第 1 列上的除第 1 行元素外也都变成 0,即化为了 $n-1$ 阶的情形.

设 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{D}$ 为对角矩阵.由于 $\boldsymbol{C}$ 可逆.$\boldsymbol{C}$ 可以写成初等矩阵的乘积.设 $\boldsymbol{C}=\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_t$,故有

$$\boldsymbol{P}_t^{\mathrm{T}}\cdots\boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_t=\boldsymbol{D}.$$

矩阵的右端乘一个初等矩阵 $\boldsymbol{P}_l$ 相当于对矩阵作一次初等列变换,在它的左端又乘以 $\boldsymbol{P}_l^{\mathrm{T}}$ 相当于对它作了一个同样类型的初等行变换.事实上,若 $\boldsymbol{P}_l=\boldsymbol{P}(i,j)$,则 $\boldsymbol{P}_l^{\mathrm{T}}=\boldsymbol{P}(i,j)$.所以对任一矩阵 $\boldsymbol{S}$,$\boldsymbol{P}_l^{\mathrm{T}}\boldsymbol{S}\boldsymbol{P}_l$ 相当于交换 $\boldsymbol{S}$ 的第 i,j 这两列后再交换所得矩阵的第 i,j 这两行.若 $\boldsymbol{P}_l=\boldsymbol{P}(i(c)),c\neq 0$,则 $\boldsymbol{P}_l^{\mathrm{T}}=\boldsymbol{P}(i(c))$,$\boldsymbol{P}_l^{\mathrm{T}}\boldsymbol{S}\boldsymbol{P}_l$ 相当于把 $\boldsymbol{S}$ 的第 i 列乘以 c 后再把所得矩阵的第 i 行乘以 c.若 $\boldsymbol{P}_l=\boldsymbol{P}(j,i(k))$,则 $\boldsymbol{P}_l^{\mathrm{T}}=\boldsymbol{P}(i,j(k))$,$\boldsymbol{P}_l^{\mathrm{T}}\boldsymbol{S}\boldsymbol{P}_l$ 相当于把 $\boldsymbol{S}$ 的第 j 列的 k 倍加到第 i 列上后再把所得矩阵的第 j 行的 k 倍加到第 i 行上,所以每次 $\boldsymbol{P}_l^{\mathrm{T}}\boldsymbol{S}\boldsymbol{P}_l$ 都是对 $\boldsymbol{S}$ 作成对初等行列变换.这表明我们作一系列的成对初等行列变换就可以把对称矩阵化成对角矩阵了.

那么又如何求出对应的可逆矩阵 $\boldsymbol{C}$ 呢?令 $\widetilde{\boldsymbol{P}}_i=\begin{pmatrix}\boldsymbol{P}_i & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{I}_n\end{pmatrix}$.由于

$$\widetilde{\boldsymbol{P}}_t^{\mathrm{T}}\cdots\widetilde{\boldsymbol{P}}_2^{\mathrm{T}}\widetilde{\boldsymbol{P}}_1^{\mathrm{T}}\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{I}\end{pmatrix}\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_t=\begin{pmatrix}\boldsymbol{P}_t^{\mathrm{T}}\cdots\boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_t\\ \boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_t\end{pmatrix}=\begin{pmatrix}\boldsymbol{D}\\ \boldsymbol{C}\end{pmatrix},$$

而每个 $\widetilde{\boldsymbol{P}}_i^{\mathrm{T}}$ 乘在矩阵的左边也就是对矩阵的前 n 行作初等矩阵 $\boldsymbol{P}_i^{\mathrm{T}}$ 对应的初等行变换.所以我们可以构造矩阵

$$\boldsymbol{R}=\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{I}\end{pmatrix},$$

即在 n 阶矩阵 $\boldsymbol{A}$ 的下方加上一个 n 阶单位矩阵. 再对矩阵 $\boldsymbol{R}$ 作成对初等行列变换. 当我们把 $\boldsymbol{R}$ 的前 n 行 $\boldsymbol{A}$ 化成对角矩阵 $\boldsymbol{D}$ 时,相应的后 n 行得到的矩阵就是所求的可逆矩阵 $\boldsymbol{C}$. 这个方法称为成对初等行列变换法.

若对称矩阵 $\boldsymbol{A}$ 合同于一个对角矩阵,称这个对角矩阵为 $\boldsymbol{A}$ 的一个合同标准形. 类似地,若二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 等价于一个只含平方项的二次型,则这个只含平方项的二次型称为二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 的一个标准形.

例 7.2 把例 7.1 中两个二次型的矩阵化成对角矩阵,并求出对应的可逆矩阵. 进一步地,把例 7.1 中的两个二次型化为标准形,并写出所用的非退化线性替换.

解 对第 1 个矩阵

$$\boldsymbol{A}_1=\begin{pmatrix}1 & -1 & \frac{3}{2}\\ -1 & 0 & -\frac{1}{2}\\ \frac{3}{2} & -\frac{1}{2} & 4\end{pmatrix},$$

$$\begin{pmatrix}1 & -1 & \frac{3}{2}\\ -1 & 0 & -\frac{1}{2}\\ \frac{3}{2} & -\frac{1}{2} & 4\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\to\begin{pmatrix}1 & 0 & 0\\ -1 & -1 & 1\\ \frac{3}{2} & 1 & \frac{7}{4}\\ 1 & 1 & -\frac{3}{2}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\to\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 1\\ 0 & 1 & \frac{7}{4}\\ 1 & 1 & -\frac{3}{2}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$\to\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 1 & \frac{11}{4}\\ 1 & 1 & -\frac{1}{2}\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}\to\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & \frac{11}{4}\\ 1 & 1 & -\frac{1}{2}\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}.$$

这便求出

$$C_1=\begin{pmatrix}1 & 1 & -\frac{1}{2}\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}.$$

且

$$C_1^{\mathrm{T}}A_1C_1=\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & \frac{11}{4}\end{pmatrix}.$$

对第 2 个矩阵

$$A_2=\begin{pmatrix}0 & 0 & \frac{1}{2} & 0\\ 0 & 0 & 0 & \frac{1}{2}\\ \frac{1}{2} & 0 & 0 & 0\\ 0 & \frac{1}{2} & 0 & 0\end{pmatrix}.$$

$$\begin{pmatrix}0 & 0 & \frac{1}{2} & 0\\ 0 & 0 & 0 & \frac{1}{2}\\ \frac{1}{2} & 0 & 0 & 0\\ 0 & \frac{1}{2} & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\to\begin{pmatrix}\frac{1}{2} & 0 & \frac{1}{2} & 0\\ 0 & 0 & 0 & \frac{1}{2}\\ \frac{1}{2} & 0 & 0 & 0\\ 0 & \frac{1}{2} & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\to\begin{pmatrix}1 & 0 & \frac{1}{2} & 0\\ 0 & 0 & 0 & \frac{1}{2}\\ \frac{1}{2} & 0 & 0 & 0\\ 0 & \frac{1}{2} & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

$$\to\begin{pmatrix}1&0&0&0\\0&0&0&\frac{1}{2}\\\frac{1}{2}&0&-\frac{1}{4}&0\\0&\frac{1}{2}&0&0\\1&0&-\frac{1}{2}&0\\0&1&0&0\\1&0&\frac{1}{2}&0\\0&0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0&0\\0&0&0&\frac{1}{2}\\0&0&-\frac{1}{4}&0\\0&\frac{1}{2}&0&0\\1&0&-\frac{1}{2}&0\\0&1&0&0\\1&0&\frac{1}{2}&0\\0&0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0&0\\0&\frac{1}{2}&0&\frac{1}{2}\\0&0&-\frac{1}{4}&0\\0&\frac{1}{2}&0&0\\1&0&-\frac{1}{2}&0\\0&1&0&0\\1&0&\frac{1}{2}&0\\0&1&0&1\end{pmatrix}$$

$$\to\begin{pmatrix}1&0&0&0\\0&1&0&\frac{1}{2}\\0&0&-\frac{1}{4}&0\\0&\frac{1}{2}&0&0\\1&0&-\frac{1}{2}&0\\0&1&0&0\\1&0&\frac{1}{2}&0\\0&1&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-\frac{1}{4}&0\\0&\frac{1}{2}&0&-\frac{1}{4}\\1&0&-\frac{1}{2}&0\\0&1&0&-\frac{1}{2}\\1&0&\frac{1}{2}&0\\0&1&0&\frac{1}{2}\end{pmatrix}\to\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-\frac{1}{4}&0\\0&0&0&-\frac{1}{4}\\1&0&-\frac{1}{2}&0\\0&1&0&-\frac{1}{2}\\1&0&\frac{1}{2}&0\\0&1&0&\frac{1}{2}\end{pmatrix}.$$

故

$$C_2=\begin{pmatrix}1&0&-\frac{1}{2}&0\\0&1&0&-\frac{1}{2}\\1&0&\frac{1}{2}&0\\0&1&0&\frac{1}{2}\end{pmatrix}.$$

且

$$C_2^{\mathrm{T}}A_2C_2=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-\frac{1}{4}&0\\0&0&0&-\frac{1}{4}\end{pmatrix}.$$

由所计算的矩阵结果可得

$f_1(x_1,x_2,x_3)$的一个标准形为 $y_1^2-y_2^2+\frac{11}{4}y_3^2$,所作的非退化线性替换为 $X=C_1Y$.而$f_2(x_1,x_2,x_3,x_4)$的一个标准形是 $y_1^2+y_2^2-\frac{1}{4}y_3^2-\frac{1}{4}y_4^2$,所作的非退化线性替换为 $X=C_2Y$.

注 与矩阵的相似标准形不同的是,一个对称矩阵的合同标准形是不唯一的,即存在不同的可逆矩阵 C_1,C_2 使得 $C_1^{\mathrm{T}}AC_1$ 和 $C_2^{\mathrm{T}}AC_2$ 是不同的对角矩阵.例如,若

$$C_1^{\mathrm{T}}AC_1=\begin{pmatrix}d_1&&&O\\&d_2&&\\&&\ddots&\\O&&&d_n\end{pmatrix}$$

为对角矩阵.令

$$C_2=C_1\begin{pmatrix}a_1&&&O\\&a_2&&\\&&\ddots&\\O&&&a_n\end{pmatrix},$$

其中每个 $a_i\neq 0,i=1,2,\cdots,n$,且 a_i 不全为±1,则 C_2 依然可逆,且

$$C_2^{\mathrm{T}}AC_2=\begin{pmatrix} a_1^2d_1 & & & O \\ & a_2^2d_2 & & \\ & & \ddots & \\ O & & & a_n^2d_n \end{pmatrix}$$

便是与 $C_1^{\mathrm{T}}AC_1$ 不同的对角矩阵.由此也可知道,数域 F 上的二次型的标准形也是不唯一的.但是标准形中系数不为 0 的平方项的个数是不变的.因为合同变换不改变矩阵的秩.而这个个数便是此二次型的矩阵的秩.今后我们也称对称矩阵 A 的秩为二次型 $X^{\mathrm{T}}AX$ 的秩.

特别地,若数域 F 为实数域$\mathbb{R}$,则定理 7.1 并不是新结果.由第六章定理 6.4 可以推出定理 7.1.事实上,这时 A 为实对称矩阵,所以存在正交矩阵 Q 使 $Q^{-1}AQ$ 为对角矩阵.Q 为正交矩阵表明 $Q^{-1}=Q^{\mathrm{T}}$,即 $Q^{\mathrm{T}}AQ$ 为对角矩阵.这便推出 A 合同于对角矩阵.所以对实对称矩阵 A 来说,不仅 A 与对角矩阵合同,同时我们可以找到正交矩阵 Q 使得 $Q^{\mathrm{T}}AQ$ 为对角矩阵.而对于实数域上的 n 元二次型 $X^{\mathrm{T}}AX$ 来说,可以找到正交矩阵 Q,使得经过非退化线性替换 $X=QY$,把二次型 $X^{\mathrm{T}}AX$ 化成标准形

$$\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2.$$

显然上面的系数 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 就是实对称矩阵 A 的全部特征值.若 Q 为正交矩阵,则非退化线性替换 $X=QY$ 称为**正交替换**(或称为**直角坐标变换**).

例 7.3 二次曲面方程为

$$x^2+4y^2+z^2-4xy-8xz-4yz-1=0.$$

作直角坐标变换,把它化为标准方程.并且指出它是什么二次曲面.

解 曲面方程左端的二次项部分

$$f(x,y,z)=x^2+4y^2+z^2-4xy-8xz-4yz$$

经过正交替换化为标准形.此二次型的矩阵为

$$A=\begin{pmatrix} 1 & -2 & -4 \\ -2 & 4 & -2 \\ -4 & -2 & 1 \end{pmatrix}.$$

要求正交矩阵 Q 使得 $Q^{\mathrm{T}}AQ$ 为对角矩阵.必须用第六章中实对称矩阵对角化的方法.

矩阵 A 的特征多项式是

$$|\lambda I-A|=\begin{vmatrix} \lambda-1 & 2 & 4 \\ 2 & \lambda-4 & 2 \\ 4 & 2 & \lambda-1 \end{vmatrix}=(\lambda-5)^2(\lambda+4).$$

因此 A 的全部特征值是 5(二重)和−4.

对于特征值 5,求出齐次线性方程组$(5I-A)X=0$ 的一个基础解系

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\-2\\0\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}1\\0\\-1\end{pmatrix}.$$

令 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1$,

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1=\begin{pmatrix}1\\0\\-1\end{pmatrix}-\frac{1}{5}\begin{pmatrix}1\\-2\\0\end{pmatrix}=\begin{pmatrix}\frac{4}{5}\\\frac{2}{5}\\-1\end{pmatrix}.$$

把 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 单位化得

$$\boldsymbol{\eta}_1=\frac{1}{|\boldsymbol{\beta}_1|}\boldsymbol{\beta}_1=\begin{pmatrix}\frac{\sqrt{5}}{5}\\-\frac{2}{5}\sqrt{5}\\0\end{pmatrix},$$

$$\boldsymbol{\eta}_2=\frac{1}{|\boldsymbol{\beta}_2|}\boldsymbol{\beta}_2=\begin{pmatrix}\frac{4}{15}\sqrt{5}\\\frac{2}{15}\sqrt{5}\\-\frac{1}{3}\sqrt{5}\end{pmatrix}.$$

对于特征值-4,求出齐次线性方程组$(-4\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的一个基础解系

$$\boldsymbol{\alpha}_3=\begin{pmatrix}2\\1\\2\end{pmatrix}.$$

把 $\boldsymbol{\alpha}_3$ 单位化得

$$\boldsymbol{\eta}_3=\frac{1}{|\boldsymbol{\alpha}_3|}\boldsymbol{\alpha}_3=\begin{pmatrix}\frac{2}{3}\\\frac{1}{3}\\\frac{2}{3}\end{pmatrix}.$$

令

$$Q=(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3)=\begin{pmatrix}\frac{\sqrt{5}}{5} & \frac{4}{15}\sqrt{5} & \frac{2}{3}\\ -\frac{2}{5}\sqrt{5} & \frac{2}{15}\sqrt{5} & \frac{1}{3}\\ 0 & -\frac{1}{3}\sqrt{5} & \frac{2}{3}\end{pmatrix},$$

则 $\boldsymbol{Q}$ 为正交矩阵,并且

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}5&0&0\\0&5&0\\0&0&-4\end{pmatrix}.$$

于是作正交替换

$$\begin{pmatrix}x\\y\\z\end{pmatrix}=\boldsymbol{Q}\begin{pmatrix}x'\\y'\\z'\end{pmatrix}, \tag{7.5}$$

可以把上面二次型 $f(x,y,z)$ 化成下面标准形

$$5x'^2+5y'^2-4z'^2,$$

因此,作直角坐标变换(7.5),上述二次曲面在新的直角坐标系中的方程为

$$5x'^2+5y'^2-4z'^2=1.$$

这是一个单叶双曲面.

§7.2 复数域上二次型的规范形

数域 F 上的二次型可以按等价关系划分为等价类.那么 F 上的二次型究竟有多少个不同的等价类?由定理 7.2,任一二次型等价于一个标准形,但二次型的标准形并不是唯一的.所以单靠标准形的概念还解决不了上面这个问题.

本节和下一节我们分别在复数域$\mathbb{C}$ 和实数域$\mathbb{R}$ 上来讨论这个问题.首先设数域 F 为复数域$\mathbb{C}$.给定$\mathbb{C}$ 上的一个二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$,其中 $\boldsymbol{A}$ 为复对称矩阵,则存在复可逆矩阵 $\boldsymbol{C}_1$ 使得

$$C_1^{\mathrm{T}}AC_1 = D = \begin{pmatrix} d_1 & & & & & & O \\ & d_2 & & & & & \\ & & \ddots & & & & \\ & & & d_r & & & \\ & & & & 0 & & \\ & O & & & & \ddots & \\ & & & & & & 0 \end{pmatrix},$$

其中 $d_i \neq 0, i=1,2,\cdots,r$，显然 D 中主对角线上非零元素个数 r 就是矩阵 A 的秩.当然也是二次型 $X^{\mathrm{T}}AX$ 的秩.

因为复数域内任意一个数都可以开平方.令

$$C_2 = \begin{pmatrix} \frac{1}{\sqrt{d_1}} & & & & & & O \\ & \frac{1}{\sqrt{d_2}} & & & & & \\ & & \ddots & & & & \\ & & & \frac{1}{\sqrt{d_r}} & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ O & & & & & & 1 \end{pmatrix},$$

则 C_2 可逆，再令 $C=C_1C_2$，则 C 可逆，且

$$C^{\mathrm{T}}AC = C_2^{\mathrm{T}}(C_1^{\mathrm{T}}AC_1)C_2 = \begin{pmatrix} \overbrace{1 & & & }^{r\text{个}} & & & O \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ O & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix} = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

这样便证明了下面的结论.

定理 7.3 设 A 为复数域 $\mathbb{C}$ 上的对称矩阵且 A 的秩 $r(A)=r$，则 A 合同于

$$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是$\mathbb{C}$ 上的秩为 r 的对称矩阵，则它们都合同于$\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$. 由合同的对称性和传递性可得 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同. 故有

定理 7.4 复数域$\mathbb{C}$ 上的对称矩阵合同当且仅当它们的秩相等.

由二次型等价与对称矩阵合同的关系，我们有下面结论.

定理 7.5 复数域$\mathbb{C}$ 上任一二次型都等价于如下一个二次型

$$y_1^2 + y_2^2 + \cdots + y_r^2, \tag{7.6}$$

其中 r 为此二次型的秩. 复数域$\mathbb{C}$ 上两个二次型等价当且仅当它们有相同的秩.

式(7.6)称为复数域上秩为 r 的二次型的规范形. 它由二次型的秩唯一决定，所以是唯一的. 因为在证明中用到对某些数做开平方运算，所以这个结论对一般数域不成立.

n 阶对称矩阵的秩有 $n+1$ 种不同的可能，所以复数域$\mathbb{C}$ 上的 n 阶对称矩阵有 $n+1$ 个不同的合同类. 这也说明复数域$\mathbb{C}$ 上的 n 元二次型有 $n+1$ 个不同的等价类.

§7.3 实数域上二次型的规范形

本节中设数域 F 为实数域$\mathbb{R}$. 对于实数域上 n 阶对称矩阵 $\boldsymbol{A}$. 存在实可逆矩阵 $\boldsymbol{C}_1$，使得

$$\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}_1 = \begin{pmatrix} d_1 & & & & & & & & \boldsymbol{O} \\ & \ddots & & & & & & & \\ & & d_p & & & & & & \\ & & & -d_{p+1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & -d_r & & & \\ & \boldsymbol{O} & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix},$$

其中 $d_i>0, i=1,2,\cdots,r, 0\leqslant p\leqslant r$，而 $r=r(\boldsymbol{A})$. 进一步地，令

$$C_2=\begin{pmatrix}\frac{1}{\sqrt{d_1}} & & & & & & & & \boldsymbol{O}\\ & \ddots & & & & & & & \\ & & \frac{1}{\sqrt{d_p}} & & & & & & \\ & & & \frac{1}{\sqrt{d_{p+1}}} & & & & & \\ & & & & \ddots & & & & \\ & \boldsymbol{O} & & & & \frac{1}{\sqrt{d_r}} & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1\end{pmatrix}$$

以及 $\boldsymbol{C}=\boldsymbol{C}_1\boldsymbol{C}_2$，则 $\boldsymbol{C}$ 可逆，且

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix}\overbrace{1 \quad \ddots \quad 1}^{p\text{ 个}} & & & & & \boldsymbol{O}\\ & & \overbrace{-1 \quad \ddots \quad -1}^{r-p\text{ 个}} & & & \\ & & & 0 & & \\ & \boldsymbol{O} & & & \ddots & \\ & & & & & 0\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_p & \boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{O} & -\boldsymbol{I}_{r-p} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}.$$

所以任一实对称矩阵 $\boldsymbol{A}$ 都合同于

$$\begin{pmatrix}\boldsymbol{I}_p & \boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{O} & -\boldsymbol{I}_{r-p} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O}\end{pmatrix},$$

其中 $r=r(\boldsymbol{A})$，p 为一不超过 r 的自然数.等价地，实数域$\mathbb{R}$ 上任一二次型 f 都等价于如下一个实二次型

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2, \tag{7.7}$$

其中 r 为该二次型 f 的秩.称(7.7)为实二次型 f 的规范形.

定理 7.6(惯性定理) 实二次型的规范形是唯一的.

证明 设 n 元实二次型 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 的秩为 r.假设它经过非退化线性替换 $\boldsymbol{X}=\boldsymbol{C}_1\boldsymbol{Y}$ 和 $\boldsymbol{X}=\boldsymbol{C}_2\boldsymbol{Z}$ 变成两个规范形

$$X^{\mathrm{T}}AX = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2, \tag{7.8}$$

$$X^{\mathrm{T}}AX = z_1^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2.$$

为证明规范形的唯一性，只需证明 $p=q$.

从(7.8)看出，经过非退化线性替换 $\boldsymbol{Y}=(\boldsymbol{C}_1^{-1}\boldsymbol{C}_2)\boldsymbol{Z}$，有

$$y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2 = z_1^2 + \cdots + z_q^2 - z_{q+1}^2 - \cdots - z_r^2. \tag{7.9}$$

令 $\boldsymbol{G}=\boldsymbol{C}_1^{-1}\boldsymbol{C}_2=(g_{ij})_{n\times n}$，则有 $\boldsymbol{Y}=\boldsymbol{GZ}$.若 $p<q$，则齐次线性方程组

$$\begin{cases} g_{11}z_1 + g_{12}z_2 + \cdots + g_{1q}z_q = 0, \\ g_{21}z_1 + g_{22}z_2 + \cdots + g_{2q}z_q = 0, \\ \cdots\cdots\cdots\cdots \\ g_{p1}z_1 + g_{p2}z_2 + \cdots + g_{pq}z_q = 0 \end{cases}$$

中方程个数小于未知量个数，有非零解.设$(k_1,k_2,\cdots,k_q)$是它的一个非零解.令

$$\begin{pmatrix} z_1 \\ \vdots \\ z_q \\ z_{q+1} \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} k_1 \\ \vdots \\ k_q \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

则可得

$$\begin{pmatrix} y_1 \\ \vdots \\ y_p \\ y_{p+1} \\ \vdots \\ y_n \end{pmatrix} = \boldsymbol{G}\begin{pmatrix} k_1 \\ \vdots \\ k_q \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ s_{p+1} \\ \vdots \\ s_n \end{pmatrix},$$

其中 $s_j=g_{j1}k_1+g_{j2}k_2+\cdots+q_{jq}k_q, j=p+1,\cdots,n$.代入(7.9)可得

$$-s_{p+1}^2 - \cdots - s_r^2 = k_1^2 + \cdots + k_q^2.$$

上式等号左端$\leqslant 0$，而等号右端>0，矛盾.因此 $p\geqslant q$.

类似地可证 $q\geqslant p$.从而 $p=q$.即实二次型的规范形是唯一的. □

定义 7.4 在秩为 r 的实二次型 f 的规范形(7.7)中，系数为 1 的平方项个数 p 称为 f 的**正惯性指数**，系数为-1 的平方项个数 $r-p$ 称为 f 的**负惯性指数**.而正惯性指数减去负惯性指数所得的差 $2p-r$ 称为 f 的**符号差**.

由上所述，实二次型的规范形被它的秩和正惯性指数决定.再利用二次型等价的对称性

和传递性立即得到如下结论.

定理 7.7 两个 n 元实二次型等价当且仅当它们的规范形相同,也当且仅当它们的秩相等并且正惯性指数也相等.

对于实对称矩阵 $\boldsymbol{A}$ 来说,$\boldsymbol{A}$ 合同于

$$\begin{pmatrix} \boldsymbol{I}_p & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{I}_{r-p} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \end{pmatrix},$$

其中 $r=r(\boldsymbol{A})$.由惯性定理可知,p 也是被 $\boldsymbol{A}$ 所唯一确定的,也称其为 $\boldsymbol{A}$ 的正惯性指数.$r-p$ 称为 $\boldsymbol{A}$ 的负惯性指数,$p-(r-p)=2p-r$ 称为 $\boldsymbol{A}$ 的符号差.而上面这个对角矩阵也称为 $\boldsymbol{A}$ 的合同规范形.由定理 7.7 立得

定理 7.8 两个 n 阶实对称矩阵合同当且仅当它们的合同规范形相同也当且仅当它们的秩相等且它们的正惯性指数也相等.

§ 7.4 正定二次型和正定矩阵

本节中出现的二次型和矩阵都是实数域上的.

定义 7.5 称 n 元实二次型 $f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$(或 n 阶实对称矩阵 $\boldsymbol{A}$)是**正定**的,如果对于 $\mathbb{R}^n$ 中任意非零列向量 $\boldsymbol{\alpha}$ 都有 $f(\boldsymbol{\alpha})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}>0$.

定理 7.9 n 元实二次型 $f(\boldsymbol{X})$ 或 n 阶实对称矩阵 $\boldsymbol{A}$ 正定当且仅当它的正惯性指数为 n.

证明 下面只对二次型的情形来证明.

必要性:设 $f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 为正定的.作非退化线性替换 $\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$ 化为规范形,即

$$f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2.$$

如果 $p<n$,则 y_n^2 的系数为 0 或-1.取

$$\boldsymbol{\beta}=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

令 $\boldsymbol{\alpha}=\boldsymbol{C}\boldsymbol{\beta}$,显然 $\boldsymbol{\alpha}\neq\boldsymbol{0}$,并且有 $f(\boldsymbol{\alpha})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=0$ 或-1,矛盾.因此 $p=n$.

充分性:设 $f(\boldsymbol{X})$ 的正惯性指数为 n,则作非退化线性替换 $\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$ 化成的规范形为

$$f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=y_1^2+y_2^2+\cdots+y_n^2.$$

对任意 $\boldsymbol{\alpha}\in\mathbb{R}^n,\boldsymbol{\alpha}\neq\mathbf{0}$,令

$$\boldsymbol{\beta}=\boldsymbol{C}^{-1}\boldsymbol{\alpha}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix},$$

则 $\boldsymbol{\beta}\neq\mathbf{0}$,且

$$f(\boldsymbol{\alpha})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=b_1^2+b_2^2+\cdots+b_n^2>0,$$

因此二次型 $f(\boldsymbol{X})$ 正定. □

由定理 7.9 可以立即得到下面的推论.

推论 7.1 n 元实二次型是正定的当且仅当它的规范形是 $y_1^2+y_2^2+\cdots+y_n^2$ 也当且仅当它的标准形中 n 个系数都是正数.

推论 7.2 n 阶实对称矩阵 $\boldsymbol{A}$ 是正定的当且仅当 $\boldsymbol{A}$ 合同于单位矩阵 $\boldsymbol{I}_n$ 也当且仅当 $\boldsymbol{A}$ 的合同标准形中主对角元全大于零,从而也当且仅当 $\boldsymbol{A}$ 的特征值全大于零.

命题 7.2 正定矩阵的行列式大于零.

证明 设 $\boldsymbol{A}$ 为 n 阶正定矩阵,则 $\boldsymbol{A}$ 与单位矩阵 $\boldsymbol{I}_n$ 合同.从而存在实可逆矩阵 $\boldsymbol{C}$ 使得

$$\boldsymbol{A}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{I}_n\boldsymbol{C}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}.$$

因此

$$|\boldsymbol{A}|=|\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}|=|\boldsymbol{C}^{\mathrm{T}}||\boldsymbol{C}|=|\boldsymbol{C}|^2>0.$$ □

注意上面命题的逆命题是不成立的.

下面从子式的角度来研究实对称矩阵正定的条件.由 §5.4 知任取矩阵 $\boldsymbol{A}$ 的 k 行 k 列.由位于这些行与列交叉位置的元素,按原来的次序组成的 k 阶行列式即为 $\boldsymbol{A}$ 的一个 k 阶子式.若取的是 $\boldsymbol{A}$ 的第 $i_1,i_2,\cdots,i_k$ 行,第 $j_1,j_2,\cdots,j_k$ 列,所得的 k 阶子式记为 $\det\left(\boldsymbol{A}_k\begin{pmatrix}i_1&i_2&\cdots&i_k\\j_1&j_2&\cdots&j_k\end{pmatrix}\right)$.

定义 7.6 设 $\boldsymbol{A}$ 是一个 n 阶矩阵.$\boldsymbol{A}$ 的一个子式称为**主子式**.如果它的行指标与列指标相同,即形如 $\det\left(\boldsymbol{A}_k\begin{pmatrix}i_1&i_2&\cdots&i_k\\i_1&i_2&\cdots&i_k\end{pmatrix}\right)$ 的子式,其中 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$.而 $\boldsymbol{A}$ 的下述主子式 $\det\left(\boldsymbol{A}_k\begin{pmatrix}1&2&\cdots&k\\1&2&\cdots&k\end{pmatrix}\right)$ 称为 $\boldsymbol{A}$ 的 k **阶顺序主子式**.

显然,n 阶矩阵的 k 阶子式共有 $\mathrm{C}_n^k\cdot\mathrm{C}_n^k=(\mathrm{C}_n^k)^2$ 个,而 k 阶主子式共有 C_n^k 个,k 阶顺序主子式只有 1 个.例如矩阵

$$\boldsymbol{A}=\begin{pmatrix}3&-1&2\\5&2&0\\-1&3&4\end{pmatrix}$$

的顺序主子式分别为

$$\det\left(A\begin{pmatrix}1\\1\end{pmatrix}\right)=|3|,\quad \det\left(A\begin{pmatrix}1&2\\1&2\end{pmatrix}\right)=\begin{vmatrix}3&-1\\5&2\end{vmatrix},\quad \det\left(A\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}\right)=|A|.$$

定理 7.10 实对称矩阵 A 正定的充分必要条件是 A 的所有顺序主子式全大于零.

证明 必要性:设 n 阶实对称矩阵 A 正定.对任意 $1\leqslant k\leqslant n$,把 A 写成分块形式

$$A=\begin{pmatrix}A_k & B_1\\ B_1^{\mathrm{T}} & B_2\end{pmatrix},$$

显然 $|A_k|$ 就是 A 的 k 阶顺序主子式.下面我们证明 A_k 为 k 阶正定矩阵.由于 A 对称,显然 A_k 也是实对称矩阵.任取 $\mathbb{R}^k$ 中一个非零列向量 γ,有

$$\gamma^{\mathrm{T}}A_k\gamma=(\gamma^{\mathrm{T}}\quad 0)\begin{pmatrix}A_k & B_1\\ B_1^{\mathrm{T}} & B_2\end{pmatrix}\begin{pmatrix}\gamma\\0\end{pmatrix}=\begin{pmatrix}\gamma\\0\end{pmatrix}^{\mathrm{T}}A\begin{pmatrix}\gamma\\0\end{pmatrix},$$

由于 $\begin{pmatrix}\gamma\\0\end{pmatrix}$ 为 $\mathbb{R}^n$ 中非零列向量,又 A 正定,所以

$$\begin{pmatrix}\gamma\\0\end{pmatrix}^{\mathrm{T}}A\begin{pmatrix}\gamma\\0\end{pmatrix}>0,$$

即 $\gamma^{\mathrm{T}}A_k\gamma>0$. 故 A_k 为 k 阶正定矩阵.由命题 7.2 知,其行列式 $|A_k|>0$.

充分性:对实对称矩阵的阶数 n 作归纳.

当 $n=1$ 时,1 阶矩阵 (a_{11}),由于 $a_{11}>0$,所以 (a_{11}) 正定.

假设对于 $n-1$ 阶实对称矩阵来说结论成立.现在来看 n 阶实对称矩阵 $A=(a_{ij})_{n\times n}$.把 A 写成分块形式

$$A=\begin{pmatrix}A_{n-1} & \alpha\\ \alpha^{\mathrm{T}} & a_{nn}\end{pmatrix},$$

其中 A_{n-1} 为 $n-1$ 阶实对称矩阵,α 为一 $n-1$ 维列向量.由于 A_{n-1} 的所有顺序主子式就是 A 的 1 阶至 $n-1$ 阶顺序主子式.由已知条件,它们均大于零,所以由归纳假设,A_{n-1} 是正定矩阵.因此有 $n-1$ 阶实可逆矩阵 C_1 使得

$$C_1^{\mathrm{T}}A_{n-1}C_1=I_{n-1}.$$

由于 A_{n-1} 可逆(正定矩阵一定可逆),令

$$U=\begin{pmatrix}I_{n-1} & -A_{n-1}^{-1}\alpha\\ 0 & 1\end{pmatrix},$$

则 U 可逆,且

$$U^{\mathrm{T}}AU=\begin{pmatrix}I_{n-1} & 0\\ -\alpha^{\mathrm{T}}A_{n-1}^{-1} & 1\end{pmatrix}\begin{pmatrix}A_{n-1} & \alpha\\ \alpha^{\mathrm{T}} & a_{nn}\end{pmatrix}\begin{pmatrix}I_{n-1} & -A_{n-1}^{-1}\alpha\\ 0 & 1\end{pmatrix},$$

$$=\begin{pmatrix} \boldsymbol{A}_{n-1} & \boldsymbol{0} \\ \boldsymbol{0} & a_{nn}-\boldsymbol{\alpha}^{\mathrm{T}}A_{n-1}^{-1}\boldsymbol{\alpha} \end{pmatrix}.$$

记 $b=a_{nn}-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha}$,上式两端取行列式得

$$|\boldsymbol{U}|^2|\boldsymbol{A}|=|\boldsymbol{A}_{n-1}|\cdot b$$

由于 $|\boldsymbol{U}|\neq 0,|\boldsymbol{A}|>0,|\boldsymbol{A}_{n-1}|>0$,所以 $b>0$.

取

$$\boldsymbol{C}=\boldsymbol{U}\begin{pmatrix} \boldsymbol{C}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{\sqrt{b}} \end{pmatrix}.$$

则 $\boldsymbol{C}$ 可逆,且

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix} \boldsymbol{C}_1^{\mathrm{T}} & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{\sqrt{b}} \end{pmatrix}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}\begin{pmatrix} \boldsymbol{C}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{\sqrt{b}} \end{pmatrix}$$

$$=\begin{pmatrix} \boldsymbol{C}_1^{\mathrm{T}} & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{\sqrt{b}} \end{pmatrix}\begin{pmatrix} \boldsymbol{A}_{n-1} & \boldsymbol{0} \\ \boldsymbol{0} & b \end{pmatrix}\begin{pmatrix} \boldsymbol{C}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \frac{1}{\sqrt{b}} \end{pmatrix}$$

$$=\begin{pmatrix} \boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{A}_{n-1}\boldsymbol{C}_1 & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{pmatrix}=\boldsymbol{I}_n.$$

这便证明 $\boldsymbol{A}$ 合同于单位矩阵 $\boldsymbol{I}_n$.从而 $\boldsymbol{A}$ 正定. □

由此定理立得

推论 7.3 实二次型正定的充分必要条件是它的矩阵的所有顺序主子式全大于零.

例 7.4 判断下述实二次型是否正定:

$$f(x_1,x_2,x_3)=x_1^2+2x_2^2+3x_3^2+4x_1x_2+2x_1x_3+2x_2x_3.$$

解 $f(x_1,x_2,x_3)$的矩阵是

$$\boldsymbol{A}=\begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

由于 $\boldsymbol{A}$ 的 2 阶顺序主子式

$$\begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix}=-2<0,$$

因此实二次型 $f(x_1,x_2,x_3)$不是正定的.

实二次型 $f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$(或实对称矩阵 $\boldsymbol{A}$)除了有正定的以外,还有其他一些类型.

（1）**半正定二次型（矩阵）**

对$\mathbb{R}^n$中任意非零列向量$\boldsymbol{\alpha}$均有$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}\geqslant 0$. 易证实二次型$f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$（或实对称矩阵$\boldsymbol{A}$）半正定当且仅当它的正惯性指数$p=$秩$r$.

（2）**负定二次型（矩阵）**

对$\mathbb{R}^n$中任意非零列向量$\boldsymbol{\alpha}$均有$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}<0$. 易证实二次型$f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$（或实对称矩阵$\boldsymbol{A}$）负定当且仅当它的负惯性指数$r-p=n$.这等价于秩$r=n$且正惯性指数$p=0$.

（3）**半负定二次型（矩阵）**

对$\mathbb{R}^n$中任意非零列向量$\boldsymbol{\alpha}$均有$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}\leqslant 0$. 易证实二次型$f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$（或实对矩阵$\boldsymbol{A}$）半负定当且仅当它的负惯性指数$r-p=r$.这等价于它的正惯性指数$p=0$.

（4）**不定二次型（矩阵）**

除正定、半正定、负定、半负定外的二次型（或实对称矩阵）.

故n元半正定二次型的规范形为$y_1^2+y_2^2+\cdots+y_r^2$，其中$0\leqslant r\leqslant n$；n元负定二次型的规范形为$-y_1^2-y_2^2-\cdots-y_n^2$；$n$元半负定二次型的规范形为$-y_1^2-y_2^2-\cdots-y_r^2$，其中$0\leqslant r\leqslant n$；而$n$元不定二次型的规范形为

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2,\quad \text{其中 } 0<p<r\leqslant n.$$

由上面定义易知实对称矩阵$\boldsymbol{A}$负定当且仅当$-\boldsymbol{A}$正定.而$\boldsymbol{A}$半负定当且仅当$-\boldsymbol{A}$半正定.

习 题 七

1. 作非退化线性替换把数域F上下列二次型化成标准形，并写出所作的非退化线性替换.

（1）$f(x_1,x_2,x_3)=x_1^2+2x_2^2+2x_1x_2-2x_1x_3$；

（2）$f(x_1,x_2,x_3)=x_1^2-x_3^2+2x_1x_2+2x_2x_3$；

（3）$f(x_1,x_2,x_3)=x_1x_2+x_1x_3+x_2x_3$；

（4）$f(x_1,x_2,x_3,x_4)=2x_1x_2-2x_3x_4$.

2. 用正交替换把下列实二次型化成标准形，并写出所作的正交替换.

（1）$f(x_1,x_2,x_3)=2x_1^2+5x_2^2+5x_3^2+4x_1x_2-4x_1x_3-8x_2x_3$；

（2）$f(x_1,x_2,x_3,x_4)=2x_1x_2-2x_3x_4$.

3. 作直角坐标变换把下述二次曲面的方程化为标准方程.并且指出它是什么二次曲面

$$2x^2+6y^2+2z^2+8xz-1=0.$$

4. 设$\boldsymbol{A}$是数域F上的对称矩阵.证明$\boldsymbol{A}=\boldsymbol{O}$当且仅当对于$F^n$中任一列向量$\boldsymbol{\alpha}$都有$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=0$.

5. 证明矩阵合同的反身性、对称性和传递性，即证明

（1）任一方阵都与自身合同；

（2）若$\boldsymbol{A}$与$\boldsymbol{B}$合同，则$\boldsymbol{B}$与$\boldsymbol{A}$合同；

(3) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同,$\boldsymbol{B}$ 与 $\boldsymbol{C}$ 合同,则 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 合同.

6. 证明秩为 r 的对称矩阵可以写成 r 个秩为 1 的对称矩阵的和.

7. 设 k 是实对称矩阵 $\boldsymbol{A}$ 的一个特征值.证明存在 $\mathbb{R}^n$ 中非零列向量 $\boldsymbol{\alpha}$ 使得 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=k\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}$.

8. 用复数域 $\mathbb{C}$ 上非退化线性替换将下列复二次型化为规范形,并求出所作的非退化线性替换

(1) $f(x_1,x_2,x_3)=-5x_1^2+3x_2^2-2x_3^2$;

(2) $f(x_1,x_2,x_3)=2x_1x_2-6x_2x_3+2x_1x_3$;

(3) $f(x_1,x_2)=(-1-\mathrm{i})x_1x_2+2\mathrm{i}x_2^2$.

9. 用实数域 $\mathbb{R}$ 上非退化线性替换将下列实二次型化为规范形,求所作的非退化线性替换并求其秩、正、负惯性指数和符号差.

(1) $f(x_1,x_2,x_3)=-5x_1^2+3x_2^2-2x_3^2$;

(2) $f(x_1,x_2,x_3)=2x_1x_2-6x_2x_3+2x_1x_3$;

(3) $f(x_1,x_2,x_3,x_4)=-(x_1-2x_2+3x_3-x_4)^2+(2x_1-x_3+3x_4)^2+(-2x_1+4x_2-6x_3+2x_4)^2$.

10. 3 阶实对称矩阵组成的集合按照合同关系分类,可以分成多少类?请在每一类里写出一个最简单的矩阵(即合同规范形).一般地,n 阶实对称矩阵组成的集合按照合同关系分类,可以分成多少类?

11. 设 $f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 是一个 n 元实二次型,且存在 $\mathbb{R}^n$ 中列向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 使得 $f(\boldsymbol{\alpha}_1)=\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}_1>0$,$f(\boldsymbol{\alpha}_2)=\boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}_2<0$. 证明存在 $\mathbb{R}^n$ 中非零列向量 $\boldsymbol{\alpha}_3$ 使得 $f(\boldsymbol{\alpha}_3)=\boldsymbol{\alpha}_3^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}_3=0$.

12. 证明与正定二次型等价的二次型依然正定,与正定矩阵合同的矩阵依然正定.

13. 证明两个半正定矩阵的和还是半正定矩阵.若其中之一为正定矩阵,则和为正定矩阵.

14. 证明正定矩阵的逆仍为正定矩阵.

15. 证明正文中所列的半正定、负定、半负定的充要条件.

16. 判断下列二次型是否正定.

(1) $f(x_1,x_2,x_3)=99x_1^2-12x_1x_2+48x_1x_3+130x_2^2-60x_2x_3+71x_3^2$;

(2) $f(x_1,x_2,x_3)=10x_1^2+8x_1x_2+24x_1x_3+2x_2^2-12x_2x_3+3x_3^2$.

17. t 取何值时,下列实二次型是正定的?

(1) $f(x_1,x_2,x_3)=x_1^2+x_2^2+5x_3^2+2tx_1x_2-2x_1x_3+4x_2x_3$;

(2) $f(x_1,x_2,x_3)=x_1^2+4x_2^2+x_3^2+2tx_1x_2+2x_2x_3$.

18. 证明实对称矩阵 $\boldsymbol{A}$ 正定的充分必要条件是有可逆实对称矩阵 $\boldsymbol{C}$,使得 $\boldsymbol{A}=\boldsymbol{C}^2$.

19. 证明实对称矩阵 $\boldsymbol{A}$ 正定的充分必要条件是它的所有主子式都大于零.从而若实对称矩阵 $\boldsymbol{A}$ 正定,则它的主对角线上元素都为正数.

20. (1) 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 为正定矩阵,则

$$|\boldsymbol{A}|\leqslant a_{11}a_{22}\cdots a_{nn}.$$

(2) 设 $\boldsymbol{T}=(t_{ij})$ 是 n 阶实矩阵.证明

$$|\boldsymbol{T}| \leqslant \prod_{i=1}^{n} \sqrt{t_{1i}^2 + t_{2i}^2 + \cdots + t_{ni}^2}.$$

21. 证明 n 阶实对称矩阵 $\boldsymbol{A}$ 负定的充分必要条件为它的偶数阶顺序主子式全为正数而奇数阶顺序主子式全为负数.

第八章　几何意义与应用专题

§8.1　矩阵与变换

本节中,我们以几何的观点来看矩阵与向量的乘法,即将矩阵视为向量空间之间的映射(变换).本节会讨论这种由矩阵定义的映射(变换)的性质,介绍$\mathbb{R}^2$与$\mathbb{R}^3$中一些特殊的矩阵变换,如:恒等变换、伸缩变换、投影变换、反射变换和旋转变换等.

进一步,我们还将从几何角度来解释本书之前学过的一些代数方面的概念与理论,如:矩阵的乘法、矩阵的特征值与特征向量、矩阵相似、正交矩阵等.

由上述讨论,读者既可体会矩阵代数的形象几何意义,也可感受代数理论作为工具是如何解决几何问题的.

8.1.1　矩阵映射与矩阵变换的定义

由第四章介绍的矩阵运算律,我们知 $m\times n$ 的实矩阵 $\boldsymbol{A}$ 与 n 维列向量 $\boldsymbol{\alpha}$ 的乘积为一个 m 维列向量,即 $\boldsymbol{A\alpha}\in\mathbb{R}^m$.很自然地,可将矩阵 $\boldsymbol{A}$ 视为向量空间$\mathbb{R}^n$到$\mathbb{R}^m$的一个映射.

定义 8.1　给定矩阵 $\boldsymbol{A}\in M_{m\times n}(\mathbb{R})$,则如下定义由 $\boldsymbol{A}$ 确定的**矩阵映射**:

$$\varphi_A:\mathbb{R}^n\longrightarrow\mathbb{R}^m$$

$$\boldsymbol{\alpha}\longmapsto\boldsymbol{A\alpha}$$

特别地,当 $m=n$ 时,φ_A 称为矩阵变换.通常可将 φ_A 简记为 $\boldsymbol{A}$.此时 $\boldsymbol{A}$ 既可表示一个映射,也可表示一个矩阵.$\boldsymbol{A\alpha}$ 既表示矩阵与向量的乘法,也表示列向量 $\boldsymbol{\alpha}$ 在 $\boldsymbol{A}$ 作用下的像.

例 8.1　令 $\boldsymbol{A}=\begin{pmatrix}0&1\\1&0\end{pmatrix}$,求下列列向量在矩阵变换 $\boldsymbol{A}$ 作用下的像.

$$\boldsymbol{e}_1=\begin{pmatrix}1\\0\end{pmatrix},\quad \boldsymbol{e}_2=\begin{pmatrix}0\\1\end{pmatrix},\quad \boldsymbol{\alpha}=\begin{pmatrix}-1\\-2\end{pmatrix}.$$

并用图像表示出来.

解 $A e_1=\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$, $A e_2=\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}$, $A\alpha=\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}-1\\-2\end{pmatrix}=\begin{pmatrix}-2\\-1\end{pmatrix}$,

如图 8.1,从图像上看,A 将上述三个向量均映为关于直线 $x_1=x_2$ 对称的向量,那么,对任意$\mathbb{R}^2$中向量 β,$A\beta$ 是否也如此呢?

例 8.2(平移变换) 设变换 $\varphi:\mathbb{R}^2\to\mathbb{R}^2$将所有点(向量)向右移 a_1,同时向上移 a_2,用向量表示为

$$\varphi(x)=\varphi\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}x_1+a_1\\x_2+a_2\end{pmatrix}=\begin{pmatrix}x_1\\x_2\end{pmatrix}+\begin{pmatrix}a_1\\a_2\end{pmatrix}=x+\alpha.$$

其中 $\alpha=\begin{pmatrix}a_1\\a_2\end{pmatrix}$,由定义 8.1 知,平移变换一般不是矩阵变换,这一点从下一小节矩阵映射的性质 1 也可看出.

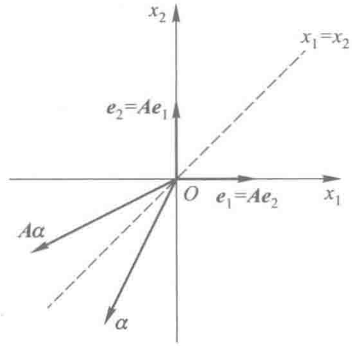


图 8.1

8.1.2 矩阵映射的性质

根据矩阵代数的运算律,我们可得矩阵映射有如下性质.

性质 1 矩阵映射把零向量映为零向量(保零元).

性质 2 矩阵映射保持向量的线性运算.即 $\forall\,\alpha,\beta\in\mathbb{R}^n$, $\forall\,k_1,k_2\in\mathbb{R}$,有

$$A(k_1\alpha+k_2\beta)=k_1(A\alpha)+k_2(A\beta).$$

反过来,若一个映射 $\sigma:\mathbb{R}^n\to\mathbb{R}^m$是保持向量线性运算的(称为线性映射),设 $e_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$, $e_2=\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}$, $\cdots$, $e_n=\begin{pmatrix}0\\0\\\vdots\\0\\1\end{pmatrix}$为$\mathbb{R}^n$的自然基,则 $\forall\,\alpha\in\mathbb{R}^n$,若 $\alpha=\begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}$,则

$$\alpha=\begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}=\sum_{i=1}^{n}a_ie_i.$$

于是有

$$\sigma(\boldsymbol{\alpha})=\sum_{i=1}^{n}a_i\sigma(\boldsymbol{e}_i)=(\sigma(\boldsymbol{e}_1),\cdots,\sigma(\boldsymbol{e}_n))\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\overset{\Delta}{=}\boldsymbol{\Phi\alpha}. \tag{8.1}$$

其中 $\boldsymbol{\Phi}$ 为一个 $m\times n$ 实矩阵.它的第 j 列就是第 j 个自然基 $\boldsymbol{e}_j$ 在 σ 作用下的像 $\sigma(\boldsymbol{e}_j)(j=1,2,\cdots,n)$.$\boldsymbol{\Phi}$ 也称为线性映射在自然基下的矩阵表示.且(8.1)式说明任何一个保持线性运算的映射均为矩阵映射.

性质 3 矩阵映射把子空间映为子空间(保子空间).

证明 设 $W\subset\mathbb{R}^n$ 为子空间,$\forall\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\in\varphi_A(W)\subset\mathbb{R}^m$,$\exists\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in\mathbb{R}^n$,使得 $\boldsymbol{\beta}_1=\boldsymbol{A\alpha}_1,\boldsymbol{\beta}_2=\boldsymbol{A\alpha}_2$.

由性质 2 知,$\forall k_1,k_2\in\mathbb{R}$,有

$$k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2=k_1(\boldsymbol{A\alpha}_1)+k_2(\boldsymbol{A\alpha}_2)=\boldsymbol{A}(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2).$$

因 W 为子空间,故 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\in W$,从而

$$k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2\in\varphi_A(W).$$

这说明 $\varphi_A(W)$ 对于加法与数乘运算是封闭的.故 $\varphi_A(W)$ 为$\mathbb{R}^m$中的子空间,也称为 W 在 $\boldsymbol{A}$ 作用下的像空间. □

一个很自然的问题是:若 $\dim W=s\leqslant n$,则 $\widetilde{W}\overset{\Delta}{=}\varphi_A(W)$ 的维数是多少呢?

设$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s\}$为 W 的一组基,则 $\forall\boldsymbol{\alpha}\in W$,$\boldsymbol{\alpha}$ 可表为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 的线性组合.由性质 2 知,$\forall\boldsymbol{\beta}\in\widetilde{W}$,存在 $\boldsymbol{\alpha}\in W$,使得 $\boldsymbol{A\alpha}\in W$,$\boldsymbol{A\alpha}$ 可表示为 $\boldsymbol{A\alpha}_1,\cdots,\boldsymbol{A\alpha}_s$ 的线性组合.即

$$\widetilde{W}=L(\boldsymbol{A\alpha}_1,\cdots,\boldsymbol{A\alpha}_s),$$

故 $\dim\widetilde{W}=r(\boldsymbol{A\alpha}_1,\cdots,\boldsymbol{A\alpha}_s)$.令 $\boldsymbol{B}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s)\in M_{n\times s}(\mathbb{R})$,即 $\boldsymbol{B}$ 是由 n 维列向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$ 组成的矩阵,且 $r(\boldsymbol{B})=s$,则由分块矩阵的运算知

$$\dim\widetilde{W}=r(\boldsymbol{A\alpha}_1,\cdots,\boldsymbol{A\alpha}_n)=r(\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times s}).$$

由 5.4.4 节矩阵秩的性质,知

$$r(\boldsymbol{A})+s-n\leqslant\dim\widetilde{W}\leqslant\min\{r(\boldsymbol{A}),s\}.$$

从而,子空间在矩阵映射 $\boldsymbol{A}$ 的作用下,只能变成维数相等或更低的子空间.特别地,在$\mathbb{R}^3$中,矩阵变换把直线变成直线或原点,把平面变成平面或直线或原点.

例 8.3 $\mathbb{R}^3$中,矩阵变换把线段变成什么图形?

解 设线段的端点为 A 与 B,则线段上的点满足下述条件:

$$\lambda\,\overrightarrow{OA}+(1-\lambda)\,\overrightarrow{OB}\quad(\forall\lambda\in[0,1]).$$

记$\overrightarrow{OA}=\boldsymbol{\alpha}$,$\overrightarrow{OB}=\boldsymbol{\beta}$,则在 $\boldsymbol{A}$ 作用下,得

$$\boldsymbol{A}(\lambda\boldsymbol{\alpha}+(1-\lambda)\boldsymbol{\beta})=\lambda\boldsymbol{A\alpha}+(1-\lambda)\boldsymbol{A\beta}\quad(\forall\lambda\in[0,1]),$$

这仍然是一条线段.故矩阵变换把线段变成线段或点(当 $\boldsymbol{A\alpha}=\boldsymbol{A\beta}$ 时,线段退化为一个点).

进一步思考:在$\mathbb{R}^3$中,矩阵变换把三角形、立方体各变成什么图形?

若 $m\neq n$，则 $\boldsymbol{A}$ 为非方阵，对应的矩阵映射是不同向量空间之间的映射.

当 $m>n$ 时，我们可通过添加全零列的方法把 $\boldsymbol{A}$ 补为 m 阶方阵. 这可将 $\boldsymbol{A}$ 的定义域从 $\mathbb{R}^n$ 延拓到 $\mathbb{R}^m$ 中，且矩阵作用的效果相同. 如

设 $\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\\a_{31}&a_{32}\end{pmatrix}$，$\boldsymbol{A}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$.

令 $\widetilde{\boldsymbol{A}}=\begin{pmatrix}a_{11}&a_{12}&0\\a_{21}&a_{22}&0\\a_{31}&a_{32}&0\end{pmatrix}$，则有

$$\widetilde{\boldsymbol{A}}\begin{pmatrix}x_1\\x_2*\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix},$$

其中 $*$ 可表示任意实数，可见 $\boldsymbol{A}$ 与 $\widetilde{\boldsymbol{A}}$ 的作用结果相同.

当 $m<n$ 时，同理也可通过添加零行的方法将 $\boldsymbol{A}$ 补成 n 阶方阵. 这可将 $\boldsymbol{A}$ 作用的值域嵌入到 $\mathbb{R}^n$ 中去. 如：

设 $\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\end{pmatrix}$，$\boldsymbol{A}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}y_1\\y_2\end{pmatrix}$.

令 $\widetilde{\boldsymbol{A}}=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\0&0&0\end{pmatrix}$，则 $\widetilde{\boldsymbol{A}}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\0\end{pmatrix}$.

综上，我们只需考虑 $m=n$ 时的情况，即只需考虑矩阵变换即可. 此时，$\boldsymbol{A}$ 为 n 阶方阵.

8.1.3 $\mathbb{R}^2$ 与 $\mathbb{R}^3$ 中几类特殊的矩阵变换

本部分我们将介绍平面 $\mathbb{R}^2$ 与几何空间 $\mathbb{R}^3$ 中的几类特殊的矩阵变换. 通过具体的几何图形，来进一步加深读者对矩阵变换的认识与理解.

1. 恒等变换

单位矩阵 $\boldsymbol{I}_2$，$\boldsymbol{I}_3$ 分别保持 $\mathbb{R}^2$ 与 $\mathbb{R}^3$ 中任何向量不变，从而保持所有几何图形不变，称为恒等变换.

反之，$\mathbb{R}^2$ 中保持所有向量不变的矩阵变换，只有 $\boldsymbol{I}_2$，证明如下：

设 $\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}$ 且 $\forall\,\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix}\in\mathbb{R}^2$，有 $\boldsymbol{A}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}x_1\\x_2\end{pmatrix}$. 特别取 $\boldsymbol{x}=\boldsymbol{e}_1=\begin{pmatrix}1\\0\end{pmatrix}$ 与 $\boldsymbol{x}=\boldsymbol{e}_2=\begin{pmatrix}0\\1\end{pmatrix}$，则

$$\left.\begin{aligned}\boldsymbol{A}\boldsymbol{e}_1&=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}a_{11}\\a_{21}\end{pmatrix}=\boldsymbol{e}_1=\begin{pmatrix}1\\0\end{pmatrix}\\\boldsymbol{A}\boldsymbol{e}_2&=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}a_{12}\\a_{22}\end{pmatrix}=\boldsymbol{e}_2=\begin{pmatrix}0\\1\end{pmatrix}\end{aligned}\right\}\Rightarrow\quad \boldsymbol{A}=(\boldsymbol{e}_1,\boldsymbol{e}_2)=\boldsymbol{I}_2.$$

同理，$\mathbb{R}^3$中保持所有向量不变的矩阵变换只有 $\boldsymbol{I}_3$.

恒等变换的概念适用于一般的$\mathbb{R}^n(n\geqslant 1)$. 只不过对$\mathbb{R}^4$以上的空间中，很难画出对应的几何图形.

2. 伸缩变换

设 $\boldsymbol{A}$ 为主对角线上为正数的二阶对角矩阵，即 $\boldsymbol{A}=\begin{pmatrix}a_1&0\\0&a_2\end{pmatrix}$，且 $a_1,a_2>0$. 则 $\boldsymbol{A}\boldsymbol{e}_1=\begin{pmatrix}a_1&0\\0&a_2\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}a_1\\0\end{pmatrix}=a_1\boldsymbol{e}_1$，$\boldsymbol{A}\boldsymbol{e}_2=\begin{pmatrix}a_1&0\\0&a_2\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\a_2\end{pmatrix}=a_2\boldsymbol{e}_2$. 从而 $\forall\,\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix}\in\mathbb{R}^2$，有

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\boldsymbol{A}(x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2)=a_1x_1\boldsymbol{e}_1+a_2x_2\boldsymbol{e}_2=\begin{pmatrix}a_1x_1\\a_2x_2\end{pmatrix}.$$

说明 $\boldsymbol{A}$ 作用在 $\boldsymbol{e}_1$ 方向上伸缩 a_1 倍（$a_1>1$ 时为伸长，$0<a_1<1$ 时为收缩，$a_1=1$ 时保持不变）. 同理 $\boldsymbol{A}$ 作用在 $\boldsymbol{e}_2$ 方向上伸缩 a_2 倍. 例如，设 $\boldsymbol{A}=\begin{pmatrix}1/2&0\\0&2\end{pmatrix}$，则 $\boldsymbol{A}$ 作用后，将平行四边形 $OABC$ 映为平行四边形 $OA'B'C'$（图 8.2）. 特别当 $a_1=a_2$ 时，为等比放缩，所得图形与原图形相似.

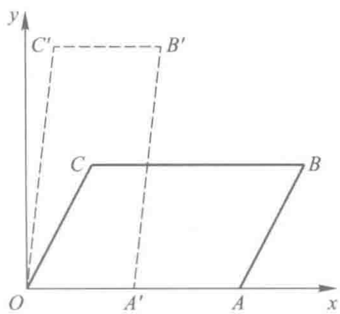


图 8.2

伸缩变换可推广到一般的 n 阶对角矩阵（主对角元为正数，$n\geqslant 2$）.

3. 反射变换

先看 2 维平面$\mathbb{R}^2$的情况. 在$\mathbb{R}^2$中有两类反射：点反射与线反射，即中心对称与轴对称.

显然任意向量 $\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix}\in\mathbb{R}^2$关于原点 O 的中心对称向量为$-\boldsymbol{x}$. 对应的变换矩阵为$-\boldsymbol{I}_2$，即 $(-\boldsymbol{I}_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}-x_1\\-x_2\end{pmatrix}=-\begin{pmatrix}x_1\\x_2\end{pmatrix}$（注：$\mathbb{R}^2$中关于原点的中心对称相当于绕原点旋转 π 的旋转变换.）.

一般地，对$\mathbb{R}^2$中任意点 $O'(a_0,b_0)$，关于 O'的中心对称如何表示呢？设 $\boldsymbol{\alpha}_0=\begin{pmatrix}a_0\\b_0\end{pmatrix}=\overrightarrow{OO'}$，$\boldsymbol{x}'$为向量 $\boldsymbol{x}$ 作变换后的坐标，由向量的平移知

$$(-\boldsymbol{I})(\boldsymbol{x}-\boldsymbol{\alpha}_0)=\boldsymbol{x}'-\boldsymbol{\alpha}_0\Rightarrow\boldsymbol{x}'=(-\boldsymbol{I})\boldsymbol{x}+2\boldsymbol{\alpha}_0. \tag{8.2}$$

但需要说明的是当 $\boldsymbol{\alpha}_0 \neq \mathbf{0}$ 时,上述点对称不是矩阵变换.下面讨论轴对称.设 $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$,则有

$$\boldsymbol{A}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}.$$

如图 8.3,$\boldsymbol{A}$ 将所有向量(图形)变为关于 x_1 轴($\boldsymbol{e}_1$)对称的向量(图形).

又设 $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,则有 $\boldsymbol{A}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$.

如图 8.4,$\boldsymbol{A}$ 将所有向量(图形)变为关于直线 $x_1 = x_2$ 对称的向量(图形).

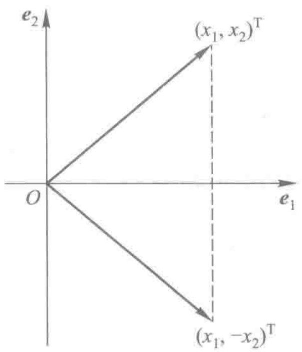


图 8.3

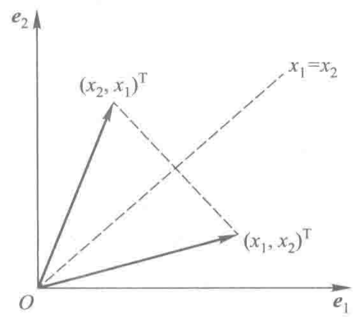


图 8.4

一般地,设对称轴为过原点的直线 L,$\boldsymbol{l}_0$ 为直线 L 的单位方向向量,对任意 $\boldsymbol{x} \in \mathbb{R}^2$,设其关于 L 的对称向量为 $\boldsymbol{x}'$,由第一章讨论的向量投影与向量内积,我们有

$$\boldsymbol{x} \cdot \boldsymbol{l}_0 = \boldsymbol{x}' \cdot \boldsymbol{l}_0. \tag{8.3}$$

另由向量加法的意义知

$$\boldsymbol{x} + \boldsymbol{x}' = k\boldsymbol{l}_0 (\exists k \in \mathbb{R}). \tag{8.4}$$

将(8.4)式代入(8.3)式得

$$\boldsymbol{x} \cdot \boldsymbol{l}_0 = (-\boldsymbol{x} + k\boldsymbol{l}_0) \cdot \boldsymbol{l}_0$$

解得 $k = 2(\boldsymbol{x} \cdot \boldsymbol{l}_0)$.再将 k 代回(8.4)式得

$$\begin{aligned} \boldsymbol{x}' &= -\boldsymbol{x} + 2(\boldsymbol{x} \cdot \boldsymbol{l}_0)\boldsymbol{l}_0 \\ &= -\boldsymbol{x} + 2(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{l}_0)\boldsymbol{l}_0 \\ &= -\boldsymbol{x} + 2(\boldsymbol{l}_0\boldsymbol{l}_0^{\mathrm{T}})\boldsymbol{x} = (-\boldsymbol{I} + 2\boldsymbol{l}_0\boldsymbol{l}_0^{\mathrm{T}})\boldsymbol{x}. \end{aligned} \tag{8.5}$$

故线反射(轴对称)对应的变换矩阵为 $\boldsymbol{R}_{l_0} = -\boldsymbol{I} + 2\boldsymbol{l}_0\boldsymbol{l}_0^{\mathrm{T}}$. 上式倒数第二个等号参见下述结论:

请验证:$\forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$(列向量),有

$$(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y})\boldsymbol{y} = (\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}})\boldsymbol{x}.$$

其中 yy^{T} 为 n 阶实对称方阵(留作习题).

今后将关于单位向量 l_0 方向的反射变换记为 R_{l_0}.

对 $L:ax+by=0$,其单位方向向量为 $l=\frac{1}{\sqrt{a^2+b^2}}(-b,a)^{\mathrm{T}}$,代入 $R_l=-I+2ll^{\mathrm{T}}$ 得

$$R_l=\begin{pmatrix}\frac{b^2-a^2}{a^2+b^2} & -\frac{2ab}{a^2+b^2}\\ \frac{-2ab}{a^2+b^2} & \frac{a^2-b^2}{a^2+b^2}\end{pmatrix}. \tag{8.6}$$

进一步,设直线 L 与 x_1 轴的夹角为 α,则由三角函数公式可得

$$R_l=\begin{pmatrix}\cos 2\alpha & \sin 2\alpha\\ \sin 2\alpha & -\cos 2\alpha\end{pmatrix}. \tag{8.7}$$

由(8.7)式易知,R_l 为正交矩阵,且 $\det R_l=-1$.

再来看几何空间$\mathbb{R}^3$的情况.

$\mathbb{R}^3$中的反射变换分三类,分别为点反射、线反射与面反射.由于(8.2)式与(8.5)式的推导并未涉及空间的维数,故$\mathbb{R}^3$中的点反射与线反射公式仍为(8.2)式与(8.5)式.

下面来看,面反射(镜面对称)变换.设 π 为$\mathbb{R}^3$中过原点的一个平面,n_0 表示其单位法向量.$\forall x\in\mathbb{R}^3$,设 x'为 x 关于平面 π 的镜面对称向量.如图 8.5,易知$-x'$恰好为 x 关于 n_0 所在直线的轴对称向量.故面反射变换的变换矩阵为

$$A=I-2n_0n_0^{\mathrm{T}}=-R_{n_0}, \tag{8.8}$$

即$\mathbb{R}^3$中的面反射等价于线反射+点反射.

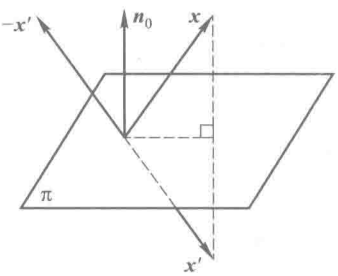


图 8.5

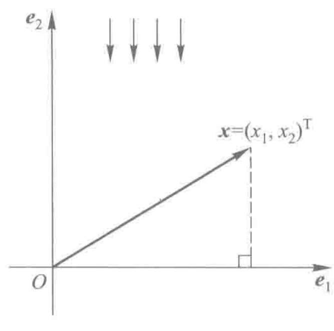


图 8.6

4. 投影变换

先看平面$\mathbb{R}^2$的情况,设 $A=\begin{pmatrix}1&0\\0&0\end{pmatrix}$,则 $A\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}x_1\\0\end{pmatrix}=x_1e_1$.

形象地看,假设有一束平行光沿着垂直于 e_1 方向(e_2 的反方向)照射.于是 xe_1 就是 x 在该光束下产生的影子(图 8.6).这称为(正交)投影变换.

同理,设 $A=\begin{pmatrix}0&0\\0&1\end{pmatrix}$,则 A 是在 e_2 方向上的(正交)投影变换.

对于$\mathbb{R}^3$中一般的情况.我们可给出如下推导.设 l 为一固定的单位向量,则对 $\forall x\in\mathbb{R}^3$,设 x'为 x 向 l 方向的(正交)投影向量,则由内积的性质知

$$x'=(x\cdot l)l=(x^{\mathrm{T}}l)l=(ll^{\mathrm{T}})x.$$

故向 l 方向(正交)投影的变换矩阵为

$$P_l=ll^{\mathrm{T}}=\begin{pmatrix}l_1\\l_2\\l_3\end{pmatrix}(l_1,l_2,l_3)=\begin{pmatrix}l_1^2&l_1l_2&l_1l_3\\l_1l_2&l_2^2&l_2l_3\\l_1l_3&l_2l_3&l_3^2\end{pmatrix},$$

称为线投影矩阵(注:$\mathbb{R}^2$中有类似结论).今后,我们记$(\alpha)_\beta$ 为 α 向 β 方向的正交投影向量.

下设 π 为过原点的平面,n 为 π 的单位法向量(如图 8.7).$\forall x\in\mathbb{R}^3$,设 x'为 x 向平面 π 正交投影的向量,则

$$x'=x-(x)_n=x-(x\cdot n)n=[I-(nn^{\mathrm{T}})]x,$$

故

$$A=I-P_n=\begin{pmatrix}1-n_1^2&-n_1n_2&-n_1n_3\\-n_1n_2&1-n_2^2&-n_2n_3\\-n_1n_3&-n_2n_3&1-n_3^2\end{pmatrix}$$

为面投影矩阵.

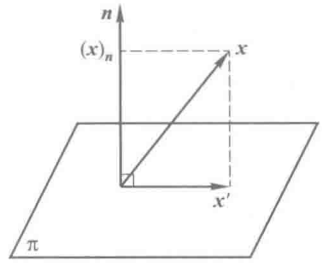


图 8.7

5. 旋转变换

在$\mathbb{R}^2$中,设 T_θ 为绕原点 O 逆时针转 θ 角的矩阵变换(图 8.8),则由

$$T_\theta e_1=\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix},\quad T_\theta e_2=\begin{pmatrix}-\sin\theta\\\cos\theta\end{pmatrix}$$

知 $T_\theta=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}\quad(0\leqslant\theta\leqslant2\pi)$.

当 $\theta=\pi$ 时,$T=\begin{pmatrix}-1&0\\0&-1\end{pmatrix}=-I$,故$\mathbb{R}^2$中关于原点的中心对称变换也为旋转变换.

在$\mathbb{R}^3$中,设 $T=\begin{pmatrix}\cos\theta&-\sin\theta&0\\\sin\theta&\cos\theta&0\\0&0&1\end{pmatrix}$,则易知 T 保持 e_3 不动,而在 e_1-e_2 平面内为逆时针旋转 θ 角.那么对于 e_1-e_2 平面之外的向量 x,在 T 的作用下,易知 x 与 Tx 的第三分量相同(设为 a_3),在平面 $x_3=a_3$ 内,令 $O'=(0,0,a_3)$,则 x 与 Tx 在平面 $x_3=a_3$ 内的投影向量夹角为 θ.故 T 的作用相当于以 e_3 为轴,将向量 x 按右手法则转动 θ

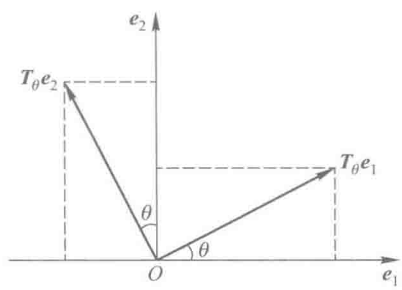


图 8.8

角(图 8.9).

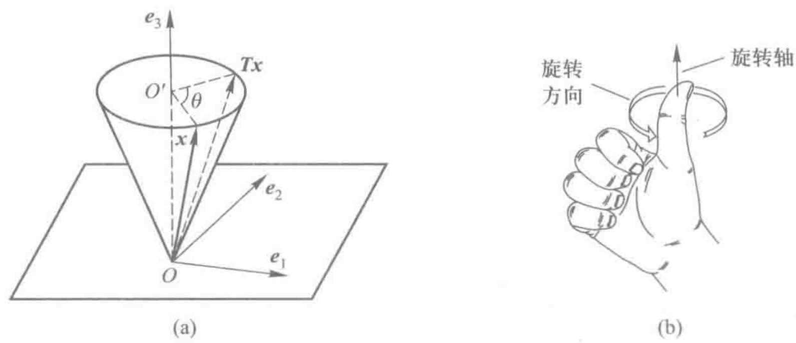


图 8.9

当 θ 跑遍 $[0,2\pi)$ 时,$\boldsymbol{Tx}$ 的运动轨迹是一个圆锥.

同理,$\boldsymbol{T}=\begin{pmatrix}\cos\theta & 0 & \sin\theta\\ 0 & 1 & 0\\ -\sin\theta & 0 & \cos\theta\end{pmatrix}$ 为 $\mathbb{R}^3$ 中绕 $\boldsymbol{e}_2$ 方向旋转 θ 角的矩阵变换,而 $\boldsymbol{T}=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 0 & \sin\theta & \cos\theta\end{pmatrix}$ 为 $\mathbb{R}^3$ 中绕 $\boldsymbol{e}_1$ 方向旋转 θ 角的矩阵变换.

一般地,设 $\boldsymbol{l}=(a,b,c)^{\mathrm{T}}$ 为给定的单位向量,令 $\boldsymbol{T}_{l,\theta}$ 为绕 $\boldsymbol{l}$ 方向旋转 θ 角的矩阵变换.那么 $\boldsymbol{T}_{l,\theta}$ 为何矩阵呢?

定理 8.1　以单位向量 $\boldsymbol{l}=(a,b,c)^{\mathrm{T}}$ 为旋转轴的方向向量,旋转角为 θ 的旋转变换的变换矩阵为

$$\boldsymbol{T}_{l,\theta}=\begin{pmatrix}a^2+(1-a^2)\cos\theta & ab(1-\cos\theta)-c\sin\theta & ac(1-\cos\theta)+b\sin\theta\\ ab(1-\cos\theta)+c\sin\theta & b^2+(1-b^2)\cos\theta & bc(1-\cos\theta)-a\sin\theta\\ ac(1-\cos\theta)-b\sin\theta & bc(1-\cos\theta)+a\sin\theta & c^2+(1-c^2)\cos\theta\end{pmatrix}.$$

定理 8.1 有多种证明方法,我们将采用一种较简单的方法,这需要用到坐标系替换与矩阵相似的结论,故我们将在 8.1.6 节给出详细证明.

8.1.4　矩阵映射的复合与矩阵乘法

我们在 4.1.2 节定义了两个矩阵的乘法,那时,我们只给出了矩阵乘法的规定,但并没有解释为什么要如此规定,诸如

- 矩阵的可乘条件($\boldsymbol{A}_{m\times n}$ 与 $\boldsymbol{B}_{s\times l}$ 可乘 $\Leftrightarrow s=n$).
- 乘积矩阵的行数与列数($\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times l}=\boldsymbol{C}_{m\times l}$).
- 乘积矩阵元素的计算公式($c_{ij}=\sum\limits_{k=1}^{n}a_{ik}b_{kj},1\leqslant i\leqslant m,1\leqslant j\leqslant l$).

对初学者来说，记住上述规定已是不容易，对这些规定的来源恐怕更是摸不着头脑.

而在8.1.1节中，对任意矩阵$\boldsymbol{A}_{m\times n}$，我们定义了矩阵映射$\varphi_A:\mathbb{R}^n\to\mathbb{R}^m$.因为$\varphi_A(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}$，故我们通常就将$\varphi_A$记为$\boldsymbol{A}$.本小节中，我们将从矩阵映射复合的角度来解释矩阵乘法.由此，以上对矩阵乘法的若干规定就变得顺理成章了.

设$\varphi_B:\mathbb{R}^l\to\mathbb{R}^n$，$\varphi_A:\mathbb{R}^n\to\mathbb{R}^m$为矩阵映射，$\forall\,\boldsymbol{x}\in\mathbb{R}^l$，$\varphi_B(\boldsymbol{x})=\boldsymbol{B}\boldsymbol{x}\in\mathbb{R}^n$.故$\varphi_B(\boldsymbol{x})$可被$\varphi_A$作用，映到$\mathbb{R}^m$中，即复合映射

$$\varphi_A(\varphi_B(\boldsymbol{x}))=\varphi_A\circ\varphi_B(\boldsymbol{x})\in\mathbb{R}^m.$$

由φ_A,φ_B均保线性运算，容易验证$\varphi_A\circ\varphi_B$仍保线性运算.由§8.1.1节中的讨论知，$\varphi_A\circ\varphi_B$必为矩阵映射.设它对应的矩阵为$\boldsymbol{C}$，则$\boldsymbol{C}\in M_{m\times l}(\mathbb{R})$.由于$\varphi_B(\boldsymbol{x})=\boldsymbol{B}\boldsymbol{x}$，$\varphi_A(\boldsymbol{y})=\boldsymbol{A}\boldsymbol{y}$，有

$$\varphi_A\circ\varphi_B(\boldsymbol{x})=\varphi_A(\varphi_B(\boldsymbol{x}))=\varphi_A(\boldsymbol{B}\boldsymbol{x})=\boldsymbol{A}(\boldsymbol{B}\boldsymbol{x})=\boldsymbol{A}\boldsymbol{B}\boldsymbol{x}.$$

上式最后一个等号由结合律得，故复合映射$\varphi_A\circ\varphi_B$对应的矩阵就为$\boldsymbol{C}=\boldsymbol{A}\boldsymbol{B}$.这就解释了矩阵可乘的条件与乘积矩阵的行数与列数为何如上规定了.

下面来看乘积矩阵的元素计算问题.还是从映射的复合来看.设$\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_l\}$为$\mathbb{R}^l$的自然基，$\{\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n'\}$为$\mathbb{R}^n$的自然基，$\{\boldsymbol{e}_1'',\cdots,\boldsymbol{e}_m''\}$为$\mathbb{R}^m$的自然基.因为$\boldsymbol{A}$的第$k$列为$\varphi_A(\boldsymbol{e}_k')=\boldsymbol{A}\boldsymbol{e}_k(1\leqslant k\leqslant n)$，$\boldsymbol{B}$的第$j$列为$\varphi_B(\boldsymbol{e}_j)=\boldsymbol{B}\boldsymbol{e}_j(1\leqslant j\leqslant l)$，而$\boldsymbol{C}=\boldsymbol{A}\boldsymbol{B}$的第$j$列应为

$$\begin{aligned}\varphi_A\circ\varphi_B(\boldsymbol{e}_j)&=\varphi_A(\boldsymbol{B}\boldsymbol{e}_j)=\varphi_A\begin{pmatrix}b_{1j}\\\vdots\\b_{nj}\end{pmatrix}=\varphi_A\Big(\sum_{k=1}^n b_{kj}\boldsymbol{e}_k'\Big)=\sum_{k=1}^n b_{kj}\varphi_A(\boldsymbol{e}_k')\\&=\sum_{k=1}^n b_{kj}\begin{pmatrix}a_{1k}\\\vdots\\a_{mk}\end{pmatrix}=\sum_{k=1}^n b_{kj}\Big(\sum_{i=1}^m a_{ik}\boldsymbol{e}_i''\Big)=\sum_{i=1}^m\Big(\sum_{k=1}^n a_{ik}b_{kj}\Big)\boldsymbol{e}_i''\\&=\begin{pmatrix}\sum\limits_{k=1}^n a_{1k}b_{kj}\\\vdots\\\sum\limits_{k=1}^n a_{mk}b_{kj}\end{pmatrix},\end{aligned}$$

从而有$c_{ij}=\sum\limits_{k=1}^n a_{ik}b_{kj}(1\leqslant i\leqslant m,1\leqslant j\leqslant l)$.

这说明矩阵乘法的诸多规定，并不是凭空定义出来的，而是由矩阵映射的复合推导出来的.

8.1.5 变换的不变量与特征值理论

我们在§6.1中引入了矩阵的特征值与特征向量，并讨论了其求法和性质.特征值与特

征向量是矩阵内在的特性指标.那么它们在矩阵变换中代表了什么含义呢?

首先来看一个 2 维的例子.

例 8.4　设 $\boldsymbol{A}=\begin{pmatrix}1&2\\2&1\end{pmatrix}$,在平面$\mathbb{R}^2$中由三点 $E(1,1)$,$F(-1,0)$,$G(1,-1)$构成的$\triangle EFG$,在 $\boldsymbol{A}$ 作用下变成什么图形?

解　由于矩阵变换将线段变为线段,故 $\boldsymbol{A}$ 作用将$\triangle EFG$ 映为$\triangle E'F'G'$,其中$\overrightarrow{OE'}=\boldsymbol{A}(\overrightarrow{OE})=\boldsymbol{A}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}3\\3\end{pmatrix}$,$\overrightarrow{OF'}=\boldsymbol{A}(\overrightarrow{OF})=\boldsymbol{A}\begin{pmatrix}-1\\0\end{pmatrix}=\begin{pmatrix}-1\\-2\end{pmatrix}$,$\overrightarrow{OG'}=\boldsymbol{A}(\overrightarrow{OG})=\boldsymbol{A}\begin{pmatrix}1\\-1\end{pmatrix}=\begin{pmatrix}-1\\1\end{pmatrix}$.容易看出向量$\begin{pmatrix}1\\1\end{pmatrix}$经过 $\boldsymbol{A}$ 作用只改变了大小,而没有改变方向;向量$\begin{pmatrix}1\\-1\end{pmatrix}$经过 $\boldsymbol{A}$ 作用后,虽改变了方向,但仍在同一直线上;但向量$\begin{pmatrix}-1\\0\end{pmatrix}$在 $\boldsymbol{A}$ 作用后,大小与方向(所在直线)均发生了变化.

进一步,向量$\begin{pmatrix}1\\1\end{pmatrix}$所在直线与向量$\begin{pmatrix}1\\-1\end{pmatrix}$所在直线上的所有向量

$$\boldsymbol{\alpha}=\begin{pmatrix}k\\k\end{pmatrix}\text{与}\boldsymbol{\beta}=\begin{pmatrix}t\\-t\end{pmatrix}\quad(k,t\in\mathbb{R})$$

在 $\boldsymbol{A}$ 作用下分别满足 $\boldsymbol{A\alpha}=3\boldsymbol{\alpha}$,$\boldsymbol{A\beta}=(-1)\boldsymbol{\beta}$.

也就是 $\boldsymbol{A}$ 在这两条直线上的作用效果分别为伸缩与反射.由§6.1 定义知,3 与−1 就是 $\boldsymbol{A}$ 的特征值,$\begin{pmatrix}1\\1\end{pmatrix}$与$\begin{pmatrix}1\\-1\end{pmatrix}$分别为属于 3 与−1 的特征向量.它们所在直线 $L_1:x_1-x_2=0$ 与 $L_2:x_1+x_2=0$ 就分别是 3 与−1 的特征空间.

一般地,若有子空间 $W\subseteq\mathbb{R}^2$满足$\forall \boldsymbol{x}\in W$,均有 $\boldsymbol{Ax}\in W$,则称 W 是矩阵变换 $\boldsymbol{A}$ 的不变子空间,$\mathbb{R}^2$中 0 维与 2 维子空间是平凡子空间.它们显然是所有矩阵变换的不变子空间.故我们只关心是否有 1 维的不变子空间.

由上讨论,特征子空间 L_1 与 L_2 就是 $\boldsymbol{A}$ 的 1 维不变子空间.由于 L_1 与 L_2 不共线,知 $\boldsymbol{v}_1=\begin{pmatrix}1\\1\end{pmatrix}$与 $\boldsymbol{v}_2=\begin{pmatrix}1\\-1\end{pmatrix}$线性无关.$\forall\boldsymbol{x}\in\mathbb{R}^2$,$\exists k_1,k_2\in\mathbb{R}$,使得 $\boldsymbol{x}=k_1\boldsymbol{v}_1+k_2\boldsymbol{v}_2$.于是 $\boldsymbol{A}$ 作用在 $\boldsymbol{x}$ 上有

$$\boldsymbol{Ax}=\boldsymbol{A}(k_1\boldsymbol{v}_1+k_2\boldsymbol{v}_2)=k_1\boldsymbol{Av}_1+k_2\boldsymbol{Av}_2=3k_1\boldsymbol{v}_1+(-1)k_2\boldsymbol{v}_2.$$

这可以简化计算,特别是在计算 $\boldsymbol{A}^k\boldsymbol{x}$ 时.

例 8.4 说明矩阵变换在特征向量的方向上较为简单,仅为伸缩与反射变换,而特征值就为伸缩的倍数.特征向量所在的直线为 1 维不变子空间.再来看几个 3 维的例子.

例 8.5　对于$\mathbb{R}^3$中的线投影变换 $\boldsymbol{P}_l=\boldsymbol{ll}^{\mathrm{T}}$,易知 $\boldsymbol{P}_l\boldsymbol{l}=(\boldsymbol{l}\cdot\boldsymbol{l}^{\mathrm{T}})\boldsymbol{l}=\boldsymbol{l}(\boldsymbol{l}^{\mathrm{T}}\boldsymbol{l})=\boldsymbol{l}$.故投影的直线 $L(\boldsymbol{l})$就是属于 $\lambda=1$ 的 1 维特征子空间.取任意与 $\boldsymbol{l}$ 正交的向量 $\boldsymbol{v}$,有

$$\boldsymbol{P}_l\boldsymbol{v}=\boldsymbol{ll}^{T}\boldsymbol{v}=\boldsymbol{0}=0\cdot\boldsymbol{v}.$$

故 $\boldsymbol{v}$ 是属于 $\lambda=0$ 的特征向量，$\lambda=0$ 的特征子空间就是 $L(\boldsymbol{l})$ 的正交补空间 $L(\boldsymbol{l})^{\perp}$，而其维数为 2. 由前讨论有，该矩阵变换保持 $L(\boldsymbol{l})$ 中的向量不动，而将 $L(\boldsymbol{l})^{\perp}$ 的所有向量变为 $\mathbf{0}$.

这与线投影的几何意义是一致的.

对于$\mathbb{R}^3$中的面投影变换，$\boldsymbol{A}=\boldsymbol{I}-\boldsymbol{n}\boldsymbol{n}^{\mathrm{T}}=\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{n}}$. 由上讨论知，$\boldsymbol{A}$ 有 2 个特征值 $\lambda_1=0$(1 重)，$\lambda_2=1$(2 重). $\lambda_1=0$ 的特征子空间为 $L(\boldsymbol{n})$，$\lambda_2=1$ 的特征子空间为 $L(\boldsymbol{n})^{\perp}=\pi$，即投影平面，从而该矩阵变换保持投影平面 π 中的向量不变. 而将 π 的法方向上的向量变为 $\mathbf{0}$. 这与面投影的几何意义是一致的.

例 8.6 对于$\mathbb{R}^3$中的线反射变换，$\boldsymbol{R}_{\boldsymbol{l}}=-\boldsymbol{I}+2\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}=-\boldsymbol{I}+2\boldsymbol{P}_{\boldsymbol{l}}$. 同理，$\boldsymbol{R}_{\boldsymbol{l}}$ 有 2 个特征值 $\lambda_1=1$(1 重)，$\lambda_2=-1$(2 重). $\lambda_1=1$ 的特征子空间为 $L(\boldsymbol{l})$，即对称轴直线. $\lambda_2=-1$ 的特征子空间为 $L(\boldsymbol{l})^{\perp}$. 说明该矩阵变换保持 $L(\boldsymbol{l})$ 上向量不变，而将所有与 $\boldsymbol{l}$ 正交的向量变为其反向量. 这与线反射变换的几何意义是一致的.

同理，对$\mathbb{R}^3$中面反射变换 $\boldsymbol{R}_{\pi}=\boldsymbol{I}-2\boldsymbol{P}_{\boldsymbol{n}}$，$L(\boldsymbol{n})$ 为 $\lambda_1=-1$ 的特征子空间，$L(\boldsymbol{n})^{\perp}=\pi$ 为 $\lambda_2=1$ 的特征子空间. 从而该矩阵变换保持 π 中向量不变，而将 π 的法方向 $\boldsymbol{n}$ 上的向量变为其反向量. 这与面反射的几何意义是一致的.

例 8.7 对$\mathbb{R}^2$中的旋转变换 $\boldsymbol{R}_{\theta}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$，求其特征多项式为

$$|\lambda\boldsymbol{I}-\boldsymbol{R}_{\theta}|=\lambda^2-2\cos\theta\lambda+1,$$

则两个特征值为 $\lambda_{1,2}=\cos\theta\pm\sqrt{1-\cos^2\theta}\,\mathrm{i}$. 当 $\theta\neq0,\pi$ 时，$\boldsymbol{R}_{\theta}$ 无实特征值，故没有实特征向量. 因此当 $\theta\neq0,\pi$ 时，$\boldsymbol{R}_{\theta}$ 在$\mathbb{R}^2$中，除 $\mathbf{0}$ 之外，没有向量是保持方向不变或方向相反的.

8.1.6 坐标系替换与矩阵相似

我们在 § 6.2 中讨论了矩阵的相似问题，并讨论了在相似意义下矩阵的对角化问题. 本节将从几何的角度来解释矩阵的相似. 具体地说：设矩阵变换 $\boldsymbol{A}$ 在新的坐标系下所对应的矩阵为 $\boldsymbol{B}$，则 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 相似，换言之，同一变换在不同坐标系下的表示矩阵是相似的. 于是矩阵的对角化问题，从几何上来看就是试图找一组基（坐标系），使得固定变换在该基（坐标系）下的表示矩阵较为简单.

坐标系的替换是处理几何问题的常用手段，如坐标系的平移. 如图 8.10，设新的原点为 O'，$\overrightarrow{OO'}=\boldsymbol{\alpha}$，$\boldsymbol{x}$ 与 $\boldsymbol{x}'$ 分别为一点在原坐标系与新坐标系下的坐标. 易验证 $\boldsymbol{x}'=\boldsymbol{x}-\boldsymbol{\alpha}$.

下面我们只讨论保持原点不变的坐标系替换. 为简单起见，我们只讨论 2 维平面$\mathbb{R}^2$的情况. 对一般的$\mathbb{R}^n$($n\geqslant3$)有相同的结论.

如图 8.11，设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 为$\mathbb{R}^2$中不共线（线性无关）的两个向量，故 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$ 构成$\mathbb{R}^2$的一组基（坐标系）. 令

$$\boldsymbol{\alpha}_1=\begin{pmatrix}p_{11}\\ p_{21}\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}p_{12}\\ p_{22}\end{pmatrix},\quad \boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)=\begin{pmatrix}p_{11} & p_{12}\\ p_{21} & p_{22}\end{pmatrix},\tag{8.9}$$

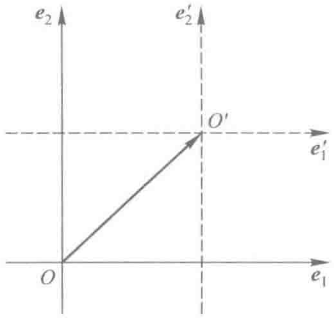


图 8.10

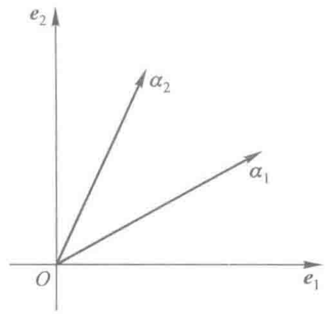


图 8.11

则 $\boldsymbol{P}$ 为可逆矩阵.对任意一点 X,设其坐标为 $\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix}$,即 $\boldsymbol{x}=(\boldsymbol{e}_1,\boldsymbol{e}_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}$,那么 X 在新的坐标系$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$下的坐标 $\boldsymbol{x}'$如何表示呢?

首先,由(8.9)式知,$\boldsymbol{P}$ 由新基的列向量组成,通常称为**过渡矩阵**.它给出了新旧两组基(坐标系)之间的联系.具体地,由

$$\boldsymbol{\alpha}_1=(\boldsymbol{e}_1,\boldsymbol{e}_2)\begin{pmatrix}p_{11}\\p_{21}\end{pmatrix},\quad \boldsymbol{\alpha}_2=(\boldsymbol{e}_1,\boldsymbol{e}_2)\begin{pmatrix}p_{12}\\p_{22}\end{pmatrix}.$$

于是有

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)=(\boldsymbol{e}_1,\boldsymbol{e}_2)\begin{pmatrix}p_{11}&p_{12}\\p_{21}&p_{22}\end{pmatrix}=(\boldsymbol{e}_1,\boldsymbol{e}_2)\boldsymbol{P}. \tag{8.10}$$

由于在新旧坐标系下表示的为同一点(向量),故有

$$(\boldsymbol{e}_1,\boldsymbol{e}_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)\begin{pmatrix}x_1'\\x_2'\end{pmatrix}.$$

将(8.10)式代入有

$$(\boldsymbol{e}_1,\boldsymbol{e}_2)\boldsymbol{x}=(\boldsymbol{e}_1,\boldsymbol{e}_2)\boldsymbol{P}\boldsymbol{x}'\Rightarrow\boldsymbol{x}=\boldsymbol{P}\boldsymbol{x}', \tag{8.11}$$

从而有

$$\boldsymbol{x}'=\boldsymbol{P}^{-1}\boldsymbol{x}. \tag{8.12}$$

(8.12)式称为**坐标替换公式**.下面讨论矩阵变换 φ_A 在坐标替换下的矩阵表示.设 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}$,在自然基$\{\boldsymbol{e}_1,\boldsymbol{e}_2\}$意义下,相当于

$$(\boldsymbol{e}_1,\boldsymbol{e}_2)\boldsymbol{A}\boldsymbol{x}=(\boldsymbol{e}_1,\boldsymbol{e}_2)\boldsymbol{y}.$$

将(8.10)式代入上式得

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{x}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)\boldsymbol{P}^{-1}\boldsymbol{y}.$$

由坐标替换公式(8.12)式,有 $\boldsymbol{x}'=\boldsymbol{P}^{-1}\boldsymbol{x},\boldsymbol{y}'=\boldsymbol{P}^{-1}\boldsymbol{y}$,代入上式有

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)\boldsymbol{P}^{-1}\boldsymbol{AP}(\boldsymbol{P}^{-1}\boldsymbol{x})=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)\boldsymbol{P}^{-1}\boldsymbol{APx}'=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)\boldsymbol{y}'.$$

上式表明,在新基$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$意义下,$\varphi_A$ 的矩阵表示为 $\boldsymbol{P}^{-1}\boldsymbol{AP}$.

这就是矩阵相似的几何意义.特征值为矩阵变换在特殊方向上(即特征子空间内)的放缩比例.同一个变换 φ_A,这样的放缩比例与坐标系的选取无关.这就是为何相似矩阵有相同特征值的原因.

例 8.8 用坐标系替换的方式重新推导$\mathbb{R}^2$内线投影变换的矩阵公式.

解 设 $\boldsymbol{l}$ 为投影直线的单位向量.选取$\{\boldsymbol{l},\boldsymbol{l}^{\perp}\}$为新的坐标系,其中 $\boldsymbol{l}=\begin{pmatrix}l_1\\l_2\end{pmatrix}$, $\boldsymbol{l}^{\perp}\triangleq\begin{pmatrix}-l_2\\l_1\end{pmatrix}$. 故过渡矩阵为

$$\boldsymbol{P}=(\boldsymbol{l},\boldsymbol{l}^{\perp})=\begin{pmatrix}l_1&-l_2\\l_2&l_1\end{pmatrix}.$$

设该变换在旧基$\{\boldsymbol{e}_1,\boldsymbol{e}_2\}$与新基$\{\boldsymbol{l},\boldsymbol{l}^{\perp}\}$下的矩阵表示分别为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$.于是有 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}$.易知 $\boldsymbol{B}=\begin{pmatrix}1&0\\0&0\end{pmatrix}$.从而有

$$\boldsymbol{A}=\boldsymbol{PBP}^{-1}=\begin{pmatrix}l_1&-l_2\\l_2&l_1\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}l_1&l_2\\-l_2&l_1\end{pmatrix}=\begin{pmatrix}l_1^2&l_1l_2\\l_1l_2&l_2^2\end{pmatrix}=\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}.$$

这与上节的结论一致.

例 8.9(一般投影变换) 设平行光束的方向为 $\boldsymbol{\beta}=\begin{pmatrix}b_1\\b_2\end{pmatrix}$,投影直线的方向向量为 $\boldsymbol{\alpha}=\begin{pmatrix}a_1\\a_2\end{pmatrix}$.设 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 不共线(也不一定正交),那么,沿 $\boldsymbol{\beta}$ 方向在 $\boldsymbol{\alpha}$ 上的投影变换所对应的矩阵是什么?

解 用坐标替换的方法来解决.以$\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$为新的坐标系.于是过渡矩阵为

$$\boldsymbol{P}=(\boldsymbol{\alpha},\boldsymbol{\beta})=\begin{pmatrix}a_1&b_1\\a_2&b_2\end{pmatrix}.$$

易知在新基下投影变换的矩阵为$\begin{pmatrix}1&0\\0&0\end{pmatrix}$,故在原基(自然基)下的矩阵为

$$\begin{aligned}\boldsymbol{P}\begin{pmatrix}1&0\\0&0\end{pmatrix}\boldsymbol{P}^{-1}&=\begin{pmatrix}a_1&b_1\\a_2&b_2\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}b_2&-b_1\\-a_2&a_1\end{pmatrix}(a_1b_2-a_2b_1)^{-1}\\&=(a_1b_2-a_2b_1)^{-1}\begin{pmatrix}a_1b_2&-a_1b_1\\a_2b_2&-a_2b_1\end{pmatrix}.\end{aligned}$$

例如:对沿着 $x_1=x_2$ 方向在 x_1 轴的投影变换(图 8.12),对应的 $\boldsymbol{\alpha}=\begin{pmatrix}1\\0\end{pmatrix}=\boldsymbol{e}_1$,$\boldsymbol{\beta}=\begin{pmatrix}1\\1\end{pmatrix}$.代入上式

得该投影的矩阵表示为$\begin{pmatrix}1 & -1\\ 0 & 0\end{pmatrix}$.

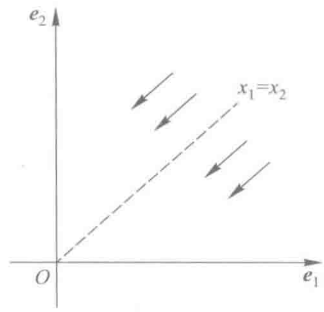


图 8.12

下面我们用坐标替换的方法证明定理 8.1,也即计算$\mathbb{R}^3$中一般旋转的矩阵.

定理 8.1 的证明:

由题设旋转轴的单位方向向量为$\boldsymbol{l}=(a,b,c)$,则$a^2+b^2+c^2=1$,记$\boldsymbol{l}=\boldsymbol{l}_1$.全体与$\boldsymbol{l}$正交的向量满足方程$ax+by+cz=0$. 易知该方程的基础解系为

$$\boldsymbol{\alpha}_2=(-b,a,0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_3=(-c,0,a)^{\mathrm{T}}.$$

对$\boldsymbol{\alpha}_2$与$\boldsymbol{\alpha}_3$做施密特正交化,得单位向量

$$\boldsymbol{l}_2=\left(\frac{-b}{\sqrt{a^2+b^2}},\frac{a}{\sqrt{a^2+b^2}},0\right)^{\mathrm{T}},\quad \boldsymbol{l}_3=\left(\frac{-ac}{\sqrt{a^2+b^2}},\frac{-bc}{\sqrt{a^2+b^2}},\sqrt{a^2+b^2}\right)^{\mathrm{T}}.$$

令 $\boldsymbol{P}=(\boldsymbol{l}_1,\boldsymbol{l}_2,\boldsymbol{l}_3)=\begin{pmatrix}a & \dfrac{-b}{\sqrt{a^2+b^2}} & \dfrac{-ac}{\sqrt{a^2+b^2}}\\ b & \dfrac{a}{\sqrt{a^2+b^2}} & \dfrac{-bc}{\sqrt{a^2+b^2}}\\ c & 0 & \sqrt{a^2+b^2}\end{pmatrix}$,则$\boldsymbol{P}$为正交矩阵,且可验证 $\det \boldsymbol{P}=1$.选取$\{0;\boldsymbol{l}_1,\boldsymbol{l}_2,\boldsymbol{l}_3\}$组成新的坐标系,于是原坐标系$\{0;\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$(自然基)到它的过渡矩阵为$\boldsymbol{P}$,且由 $\det \boldsymbol{P}=1$ 知,新旧坐标系中的右手法则保持不变.

因为在新坐标系中,绕$\boldsymbol{l}_1$方向按右手法则转θ角的变换矩阵为

$$\boldsymbol{B}=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 0 & \sin\theta & \cos\theta\end{pmatrix},$$

所以在原坐标系中,该旋转的矩阵为

$$\boldsymbol{T}_{l\theta}=\boldsymbol{PBP}^{-1}=\boldsymbol{PBP}^{\mathrm{T}}$$

$$=\begin{pmatrix}a & \dfrac{-b}{\sqrt{a^2+b^2}} & \dfrac{-ac}{\sqrt{a^2+b^2}}\\ b & \dfrac{a}{\sqrt{a^2+b^2}} & \dfrac{-bc}{\sqrt{a^2+b^2}}\\ c & 0 & \sqrt{a^2+b^2}\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 0 & \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}a & b & c\\ \dfrac{-b}{\sqrt{a^2+b^2}} & \dfrac{a}{\sqrt{a^2+b^2}} & 0\\ \dfrac{-ac}{\sqrt{a^2+b^2}} & \dfrac{-bc}{\sqrt{a^2+b^2}} & \sqrt{a^2+b^2}\end{pmatrix}$$

$$=\begin{pmatrix}a^2+(1-a^2)\cos\theta & ab(1-\cos\theta)-c\sin\theta & ac(1-\cos\theta)+b\sin\theta\\ ab(1-\cos\theta)+c\sin\theta & b^2+(1-b^2)\cos\theta & bc(1-\cos\theta)-a\sin\theta\\ ac(1-\cos\theta)-b\sin\theta & bc(1-\cos\theta)+a\sin\theta & c^2+(1-c^2)\cos\theta\end{pmatrix}\qquad\square$$

8.1.7 正交变换

在几何问题中，常常需要求保持几何图形形状的变换；在物理问题中，有时需要描述物体保持形状的运动过程（简称刚体运动）.本节就来讨论这种保持形状的矩阵变换.因为与正交矩阵的关系，我们称这样的矩阵变换为正交变换.

首先，来看何谓“保形状”？形状实际上与一个向量的长度及任意两个向量的夹角密切相关，而我们在 § 5.6 中已讨论过，长度与夹角均可由内积来定义.因此，我们可用如下“保内积”的条件来刻画矩阵变换.

定义 8.2 在$\mathbb{R}^n$中，设$(\boldsymbol{\alpha},\boldsymbol{\beta})$表示标准内积，$\varphi_A$ 为一个矩阵变换，若$\forall\,\boldsymbol{\alpha},\boldsymbol{\beta}$，均有

$$(\varphi_A(\boldsymbol{\alpha}),\varphi_A(\boldsymbol{B}))=(\boldsymbol{A\alpha},\boldsymbol{A\beta})=(\boldsymbol{\alpha},\boldsymbol{\beta}),\tag{8.13}$$

则称 φ_A 为正交变换.

注 ① 正交变换的定义可适用于任何欧氏空间.

② (8.13)式可称为保内积.

③ 由 § 5.6 讨论知，正交变换保持长度、保持夹角（统称为保度量），即 $|\boldsymbol{A\alpha}|=|\boldsymbol{\alpha}|$，$\langle\boldsymbol{A\alpha},\boldsymbol{A\beta}\rangle=\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle$ $(\forall\,\boldsymbol{\alpha},\boldsymbol{\beta})$.

④ 平移变换是保持形状的，但它不是矩阵变换，故不是正交变换.另一方面，由平移变换移动了原点，因而零向量的长度 0 变成了正数，不再保持长度.这也说明平移不是正交变换.

定理 8.2 在$\mathbb{R}^n$中，φ_A 为矩阵变换，则下述条件等价

(i) φ_A 为正交变换（保内积）；

(ii) $|\boldsymbol{A\alpha}|=|\boldsymbol{\alpha}|\,(\forall\,\boldsymbol{\alpha}\in\mathbb{R}^n)$（保长度）； (8.14)

(iii) $\boldsymbol{A}$ 为正交矩阵.

证明 (i)⇒(ii) 注③已说明.

(ii)⇒(i) $\forall\,\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{R}^n$，由(8.14)式知

$$|\boldsymbol{A}(\boldsymbol{\alpha}+\boldsymbol{\beta})|=|\boldsymbol{\alpha}+\boldsymbol{\beta}|,$$

而

$$\begin{aligned}|\boldsymbol{A}(\boldsymbol{\alpha}+\boldsymbol{\beta})|^2&=(\boldsymbol{A}(\boldsymbol{\alpha}+\boldsymbol{\beta}),\boldsymbol{A}(\boldsymbol{\alpha}+\boldsymbol{\beta}))\\&=(\boldsymbol{A\alpha}+\boldsymbol{A\beta},\boldsymbol{A\alpha}+\boldsymbol{A\beta})\\&=|\boldsymbol{A\alpha}|^2+|\boldsymbol{A\beta}|^2+2(\boldsymbol{A\alpha},\boldsymbol{A\beta}),\end{aligned}$$

$$|\boldsymbol{\alpha}+\boldsymbol{\beta}|^2=(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}+\boldsymbol{\beta})=|\boldsymbol{\alpha}|^2+|\boldsymbol{\beta}|^2+2(\boldsymbol{\alpha},\boldsymbol{\beta}).$$

对比上两式，得

$$(\boldsymbol{A\alpha},\boldsymbol{A\beta})=(\boldsymbol{\alpha},\boldsymbol{\beta}),$$

故 φ_A 为正交变换.

(i)⇒(iii) 取 $\boldsymbol{\alpha}=\boldsymbol{e}_i,\boldsymbol{\beta}=\boldsymbol{e}_j$，则

$$(\boldsymbol{A}\boldsymbol{e}_i,\boldsymbol{A}\boldsymbol{e}_j)=(\boldsymbol{e}_i,\boldsymbol{e}_j)=\delta_{ij}=\begin{cases}1, & \text{当 } i=j \text{ 时},\\ 0, & \text{当 } i\neq j \text{ 时}.\end{cases}$$

而 $\boldsymbol{A}\boldsymbol{e}_i$ 与 $\boldsymbol{A}\boldsymbol{e}_j$ 分别为 $\boldsymbol{A}$ 的第 i 列与第 j 列，这说明 $\boldsymbol{A}$ 的列构成$\mathbb{R}^n$的一组标准正交基，从而 $\boldsymbol{A}$ 为正交矩阵.

(iii)⇒(i) 设 $\boldsymbol{\alpha}=\sum\limits_{i=1}^{n}a_i\boldsymbol{e}_i,\boldsymbol{\beta}=\sum\limits_{j=1}^{n}b_j\boldsymbol{e}_j,\boldsymbol{A}\boldsymbol{e}_i$ 为 $\boldsymbol{A}$ 的第 i 列，由 $\boldsymbol{A}$ 为正交矩阵.有 $(\boldsymbol{A}\boldsymbol{e}_i,\boldsymbol{A}\boldsymbol{e}_j)=\delta_{ij}$，则

$$\begin{aligned}(\boldsymbol{A}\boldsymbol{\alpha},\boldsymbol{A}\boldsymbol{\beta})&=\left(\boldsymbol{A}\left(\sum_{i=1}^{n}a_i\boldsymbol{e}_i\right),\boldsymbol{A}\left(\sum_{j=1}^{n}b_j\boldsymbol{e}_j\right)\right)\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}a_ib_j(\boldsymbol{A}\boldsymbol{e}_i,\boldsymbol{A}\boldsymbol{e}_j)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_ib_j\delta_{ij}\\&=\sum_{i=1}^{n}a_ib_i=(\boldsymbol{\alpha},\boldsymbol{\beta}),\end{aligned}$$

故 $\varphi_{\boldsymbol{A}}$ 为正交变换.

下面来看$\mathbb{R}^2$与$\mathbb{R}^3$中，哪些变换是正交变换.

例 8.10　对$\mathbb{R}^2$中的旋转变换，$\boldsymbol{R}_\theta=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$，易知，对 $\forall\,\theta\in[0,2\pi)$，$\boldsymbol{\alpha}_1=\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$，$\boldsymbol{\alpha}_2=\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}$，总满足 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=\delta_{ij}(1\leqslant i,j\leqslant 2)$，故 $\boldsymbol{R}_\theta$ 为正交矩阵，且 $\det\boldsymbol{R}_\theta=1$，所以旋转变换为正交变换.

例 8.11　对$\mathbb{R}^n$中的线反射变换 $\boldsymbol{R}_l=-\boldsymbol{I}+2\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}$（简记为 $\boldsymbol{R}$），其中 $\boldsymbol{l}$ 为反射轴方向的单位向量.由于

$$\begin{aligned}\boldsymbol{R}\boldsymbol{R}^{\mathrm{T}}&=(-\boldsymbol{I}+2\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}})(-\boldsymbol{I}+2\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}})^{\mathrm{T}}=(-\boldsymbol{I}+2\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}})(-\boldsymbol{I}+2\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}})\\&=\boldsymbol{I}^2-4\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}+4(\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}})(\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}})\\&=\boldsymbol{I}-4\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}+4\boldsymbol{l}(\boldsymbol{l}^{\mathrm{T}}\boldsymbol{l})\boldsymbol{l}^{\mathrm{T}}\\&=\boldsymbol{I}-4\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}+4\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}=\boldsymbol{I}.\end{aligned}$$

其中 $\boldsymbol{l}^{\mathrm{T}}\boldsymbol{l}=(\boldsymbol{l},\boldsymbol{l})=1$.故 $\boldsymbol{R}_l$ 为正交矩阵.线反射变换为正交变换.

进一步，将 $\boldsymbol{l}$ 扩充为$\mathbb{R}^n$中的一组标准正交基 $\{\boldsymbol{l}_1,\boldsymbol{l}_2,\cdots,\boldsymbol{l}_n\}$ $(\boldsymbol{l}_1=\boldsymbol{l})$，则

$$\boldsymbol{R}\boldsymbol{l}_i=-\boldsymbol{l}_i+2\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}\boldsymbol{l}_i=-\boldsymbol{l}_i\quad(i=2,3,\cdots,n),$$

$$\boldsymbol{R}\boldsymbol{l}=-\boldsymbol{l}+2\boldsymbol{l}\boldsymbol{l}^{\mathrm{T}}\boldsymbol{l}=-\boldsymbol{l}+2\boldsymbol{l}=\boldsymbol{l},$$

故 ±1 是 $\boldsymbol{R}$ 的特征值，$\boldsymbol{l},\boldsymbol{l}_2,\cdots,\boldsymbol{l}_n$ 为 $\boldsymbol{R}$ 的特征向量.

由于 $\boldsymbol{l},\boldsymbol{l}_2,\cdots,\boldsymbol{l}_n$ 线性无关，故 $\boldsymbol{R}$ 可对角化，即

$$P^{\mathrm{T}}RP=\begin{pmatrix}1&&&\\&-1&&\\&&\ddots&\\&&&-1\end{pmatrix}\ (n-1\text{ 个}),$$

其中 $P=(l_1,l_2,\cdots,l_n)$ 为正交阵，

$$\det R=(-1)^{n-1}.$$

例 8.12 对 $\mathbb{R}^3$ 中的面反射变换，$R_\pi=I-2nn^{\mathrm{T}}$，其中 n 为反射平面 π 的单位法向量. 由于 $R_\pi=-R_n$，由上例可知，R_π 也为正交变换，且将 n 扩充为 $\mathbb{R}^3$ 的一组标准正交基 $\{n=n_1,n_2,n_3\}$ 后，R_π 可正交对角化，即 $Q^{\mathrm{T}}R_\pi Q=\begin{pmatrix}-1&&\\&1&\\&&1\end{pmatrix}$，其中 $Q=(n_1,n_2,n_3)$ 为正交矩阵，且 $\det R_\pi=-1$.

例 8.13 对 $\mathbb{R}^3$ 中的旋转变换，$T_{l,\theta}$（绕单位向量 $l=(a,b,c)^{\mathrm{T}}$，按右手法则逆时针转 θ 角）. 由定理 8.1 知

$$T=T_{l,\theta}=\begin{pmatrix}a^2+(1-a^2)\cos\theta & ab(1-\cos\theta)-c\sin\theta & ac(1-\cos\theta)+b\sin\theta\\ ab(1-\cos\theta)+c\sin\theta & b^2+(1-b^2)\cos\theta & bc(1-\cos\theta)-a\sin\theta\\ ac(1-\cos\theta)-b\sin\theta & bc(1-\cos\theta)+a\sin\theta & c^2+(1-c^2)\cos\theta\end{pmatrix}.$$

从代数角度可验证 T 为正交矩阵，但比较繁琐，从几何意义易知，旋转变换是保长度的，故是正交变换.

另外可直接验证 $\det T=1$（比较繁琐），也可利用 8.1.3 中 $T=T_5T_4\cdots T_1$ 的关系，易验证 $\det T_i=1(i=1,2,\cdots,5)$，故 $\det T=\prod\limits_{i=1}^{5}\det T_i=1$.

设 $\lambda_1,\lambda_2,\lambda_3$ 为 T 的三个复特征值，则

$$\lambda_1+\lambda_2+\lambda_3=\mathrm{Tr}(T)=1+2\cos\theta,$$
$$\lambda_1\lambda_2\lambda_3=\det T=1.$$

由于

$$|I-T|=|TT^{\mathrm{T}}-T|=|T||T^{\mathrm{T}}-I|=|T||T-I|=|T-I|=-|I-T|,$$

则必有 $|I-T|=0$，即 1 为 T 的特征值.

不妨设 $\lambda_1=1$，故有

$$\begin{cases}\lambda_2+\lambda_3=2\cos\theta,\\ \lambda_2\lambda_3=1.\end{cases}$$

解得 $\lambda_2=\cos\theta+\mathrm{i}\sin\theta$，$\lambda_3=\cos\theta-\mathrm{i}\sin\theta$，即 T 的一对共轭复特征值的辐角恰好为旋转变换的旋转角. v_1 为属于 1 的一个特征向量. 由几何重数≤代数重数知，属于 $\lambda_1=1$ 的特征子空间为 1 维的，就是 $L(v_1)$，由于 $Tv_1=v_1$，即 v_1 在绕轴转动 θ 角后，仍为 v_1. 当 $\theta\neq0$ 时，$L(v_1)$ 就是一

旋转轴.这就是 $T_{l,\theta}$ 特征值与特征向量的几何意义.

本小节最后,我们来证明 $\mathbb{R}^2$ 与 $\mathbb{R}^3$ 中的任意正交变换只能是旋转变换、反射变换或前两者的复合.对正交矩阵而言,$A^{\mathrm{T}}A=I\Rightarrow \det A=\pm 1$.

对 $\det A=1$ 的正交矩阵 A 所对应的变换称为第一类正交变换.

对 $\det A=-1$ 的正交矩阵 A 所对应的变换称为第二类正交变换.

定理 8.3(正交变换基本定理)　在 $\mathbb{R}^2$ 与 $\mathbb{R}^3$ 中,第一类正交变换必为旋转变换;第二类正交变换必为反射变换,或某个反射变换与某个旋转变换的复合.

证明　先证 $\mathbb{R}^2$ 的情况.设 A 为任意 2 阶正交矩阵,记 $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$,由 $A^{\mathrm{T}}A=I$ 知

$$a_{11}^2+a_{21}^2=1=a_{12}^2+a_{22}^2\Rightarrow \text{可设 } a_{11}=\cos\theta_1, a_{21}=\sin\theta_1, a_{12}=\cos\theta_2, a_{22}=\sin\theta_2.$$

又由于 $a_{11}a_{12}+a_{21}a_{22}=0,\theta_2=\theta_1\pm\dfrac{\pi}{2}$

$$\Rightarrow A=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{或 } A=\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

当 $\det A=1$ 时,只有 $A=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}=T_\theta$,为旋转变换;

当 $\det A=-1$ 时,只有 $A=\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}=T_\theta\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$,为反射变换与旋转变换的复合.

再证 $\mathbb{R}^3$ 的情况.设 A 为 3 阶正交矩阵且 $\det A=1$,则

$$|I-A|=|A^{\mathrm{T}}A-A|=|A||A^{\mathrm{T}}-I|=|A-I|=-|I-A|,$$

从而有 $|I-A|=0$,说明 1 必为 A 的特征值.

设 v_1 为属于特征值 1 的一个特征向量,且 $|v_1|=1$,将 v_1 扩充成为 $\mathbb{R}^3$ 的一组标准正交基 $\{v_1,v_2,v_3\}$,由于

$$(v_1,Av_2)=(Av_1,Av_2)=(v_1,v_2)=0,$$

知 $Av_2\in L(v_2,v_3)=L(v_1)^{\perp}$.

同理 $Av_3\in L(v_2,v_3)=L(v_1)^{\perp}$,即 $\forall \alpha\in L(v_2,v_3)$,有 $A\alpha\in L(v_2,v_3)$.

故 φ_A 在坐标系 $\{v_1,v_2,v_3\}$ 下的矩阵表示为

$$B=\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A_1 \end{pmatrix}=P^{-1}AP=P^{\mathrm{T}}AP,$$

其中 $P=(v_1,v_2,v_3)$ 为正交矩阵.

由 $B^{\mathrm{T}}B=P^{\mathrm{T}}A^{\mathrm{T}}PP^{\mathrm{T}}AP=I$,知 B 也为正交矩阵,且 $\det B=\det A=1$.这相当于 A_1 为 2 阶正交矩阵且 $\det A_1=1$.由本定理上半部分可知,A_1 必为如下形式:

$$A_1=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},\quad \theta\in[0,2\pi).$$

故 $\boldsymbol{B}=\begin{pmatrix}1&0&0\\0&\cos\theta&-\sin\theta\\0&\sin\theta&\cos\theta\end{pmatrix}$.从而 φ_A 相当于绕 $\boldsymbol{v}_1$ 方向旋转 θ 角.

对第二类正交变换 $\boldsymbol{A}$,有 $\boldsymbol{B}=\boldsymbol{A}\begin{pmatrix}-1&&\\&1&\\&&1\end{pmatrix}$ 为第一类正交变换,从而 $\boldsymbol{A}=\boldsymbol{B}\begin{pmatrix}-1&&\\&1&\\&&1\end{pmatrix}$ 为反射变换与旋转变换的复合.

§8.2　行列式的几何意义

8.2.1　二阶、三阶行列式的几何意义

在本节中,我们证明行列式的几何解释.

定理 8.4　若 $\boldsymbol{A}$ 是一个 2×2 矩阵,则由 $\boldsymbol{A}$ 的列确定的平行四边形的面积为 $|\det \boldsymbol{A}|$;若 $\boldsymbol{A}$ 是一个 3×3 矩阵,则由 $\boldsymbol{A}$ 的列确定的平行六面体的体积为 $|\det \boldsymbol{A}|$.

若 $\boldsymbol{A}$ 为 2 阶对角矩阵,定理显然成立

$$\left|\det\begin{pmatrix}a&0\\0&d\end{pmatrix}\right|=|ad|=\text{矩形的面积}.$$

如图 8.13.

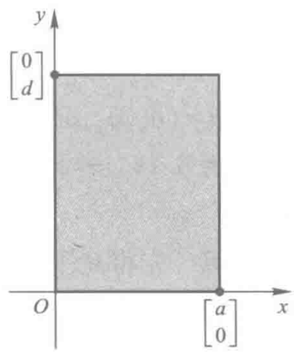


图 8.13

若 $\boldsymbol{A}$ 不为对角情形,只要证 $\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2)$ 能变换成一个对角矩阵,同时既不改变对应的平行四边形的面积,又不改变 $|\det \boldsymbol{A}|$.

对于任意矩阵 $\boldsymbol{A}$.通过交换两列或一列的倍数加到另一列上,可以使 $\boldsymbol{A}$ 变成对角矩阵,并且这两种变换都不改变行列式的值的绝对值.由于交换两列不改变对应的平行四边形的面积.因此,我们只需证明倍加变换也不改变对应的平行四边形的面积即可.先证明下述结论:

引理　设 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 为非零向量,则对任意数 c,由 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 确定的平行四边形的面积等于由 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2+c\boldsymbol{a}_1$ 确定的平行四边形的面积.

证明　不失一般性,我们假设 $\boldsymbol{a}_2$ 不是 $\boldsymbol{a}_1$ 的倍数,否则这两个平行四边形将退化成面积为 0 的情形.若 L 是通过 $\boldsymbol{0}$ 和 $\boldsymbol{a}_1$ 的直线,则 $\boldsymbol{a}_2+L$ 是通过 $\boldsymbol{a}_2$ 且平行于 L 的直线,$\boldsymbol{a}_2+c\boldsymbol{a}_1$ 在此直

线上.点 $\boldsymbol{a}_2$ 和 $\boldsymbol{a}_2+c\boldsymbol{a}_1$ 到直线 L 有相同的垂直距离,则两个平行四边形有相同的底边、相同的高,因此,它们具有相同的面积.见图 8.14.

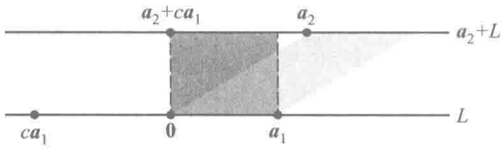


图 8.14

□

类似可证明$\mathbb{R}^3$的情形(定理对 3×3 的对角矩阵成立.见图 8.15.任意一个 3×3 矩阵 $\boldsymbol{A}$ 均可用不改变$|\det \boldsymbol{A}|$的列变换成对角矩阵,因此只要证明这些变换不改变 $\boldsymbol{A}$ 的列确定的平行六面体的体积即可.).

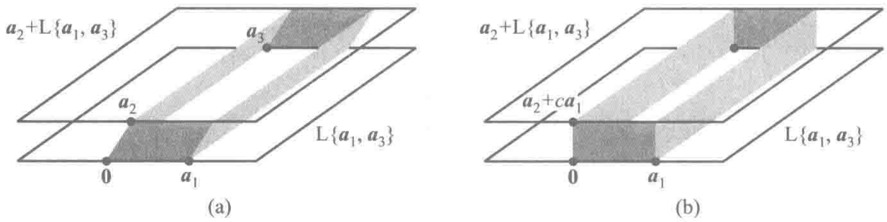


图 8.15

另外,在 1.2.3 节中讨论混合积时已经提过,以三个非零向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 为棱作一个平行六面体,则它的体积为

$$V=|(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}|=|(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})|$$
$$=\left\|\begin{matrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{matrix}\right\|.$$

其中 $\boldsymbol{a}=(a_x,a_y,a_z)^{\mathrm{T}},\boldsymbol{b}=(b_x,b_y,b_z)^{\mathrm{T}},\boldsymbol{c}=(c_x,c_y,c_z)^{\mathrm{T}}$.

例 8.14 计算由点(-2,-2),(0,3),(6,4)和(4,-1)确定的平行四边形的面积.见图 8.16.

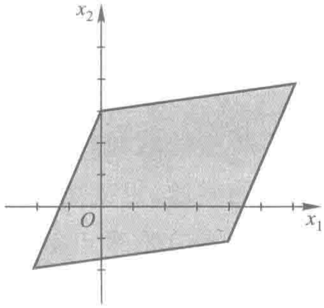
解 先将此平行四边形平移到使原点作为其一顶点的情形.

例如,将每个顶点坐标加上向量(2,2).

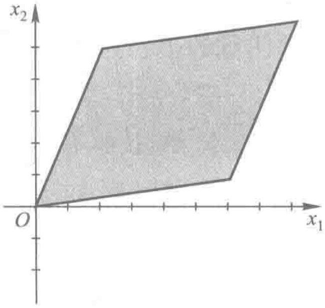
这样,新的平行四边形与原平行四边形面积相同,其顶点为(0,0),(2,5),(8,6)和(6,1).此平行四边形由 $\boldsymbol{A}=\begin{pmatrix} 2 & 6 \\ 5 & 1 \end{pmatrix}$的列所确定.由于

$$|\det \boldsymbol{A}|=|-28|=28,$$

故所求平行四边形的面积为 28.



(a)



(b)

图 8.16

8.2.2 矩阵变换与行列式

行列式可用于描述$\mathbb{R}^2$和$\mathbb{R}^3$中矩阵变换的一个重要的几何性质.若 T 是一个矩阵变换,S 是 T 的定义域内的一个集合,用 $T(S)$表示 S 中点的像集.我们关心这样的问题:集合 S 的面积(体积)与 $T(S)$的面积(体积)相比,有什么变化? 当 S 是一个边界为平行四边形的区域时,我们就用 S 表示一个平行四边形.

定理 8.5 设 $T:\mathbb{R}^2\to\mathbb{R}^2$是由一个 2×2 矩阵 $\boldsymbol{A}$ 确定的矩阵变换,若 S 是$\mathbb{R}^2$中一个平行四边形,则

$$T(S)\text{ 的面积} = |\det \boldsymbol{A}|\cdot S\text{ 的面积}. \tag{8.15}$$

若 T 是一个由 3×3 矩阵 $\boldsymbol{A}$ 确定的矩阵变换,而 S 是$\mathbb{R}^3$中的一个平行六面体,则

$$T(S)\text{ 的体积} = |\det \boldsymbol{A}|\cdot \{S\text{ 的体积}\}. \tag{8.16}$$

证明 考虑 2×2 矩阵的情形,$\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2)$.一个顶点在$\mathbb{R}^2$中原点的平行四边形,若它由向量 $\boldsymbol{b}_1,\boldsymbol{b}_2$ 确定,则有 $S=\{s_1\boldsymbol{b}_1+s_2\boldsymbol{b}_2,0\leqslant s_1\leqslant 1,0\leqslant s_2\leqslant 1\}$.$S$ 在 T 下的像由以下形式的点组成:

$$T(s_1\boldsymbol{b}_1+s_2\boldsymbol{b}_2)=s_1T(\boldsymbol{b}_1)+s_2T(\boldsymbol{b}_2)=s_1\boldsymbol{A}\boldsymbol{b}_1+s_2\boldsymbol{A}\boldsymbol{b}_2,$$

从而 $T(S)$是由矩阵$(\boldsymbol{A}\boldsymbol{b}_1,\boldsymbol{A}\boldsymbol{b}_2)$的列确定的平行四边形.

记 $\boldsymbol{A}\boldsymbol{B}=(\boldsymbol{A}\boldsymbol{b}_1,\boldsymbol{A}\boldsymbol{b}_2)$,其中矩阵 $\boldsymbol{B}=(\boldsymbol{b}_1,\boldsymbol{b}_2)$.因此,

$T(S)$的面积$=|\det \boldsymbol{A}\boldsymbol{B}|=|\det \boldsymbol{A}|\cdot|\det \boldsymbol{B}|=|\det \boldsymbol{A}|\cdot\{S$ 的面积$\}$.

对于平行四边形没有顶点在原点的情形,只需将平行四边形进行平移即可.即对任一个具有 $\boldsymbol{P}+S$ 的平行四边形,S 是一个顶点在原点的平行四边形,$\boldsymbol{P}$ 为平移向量则

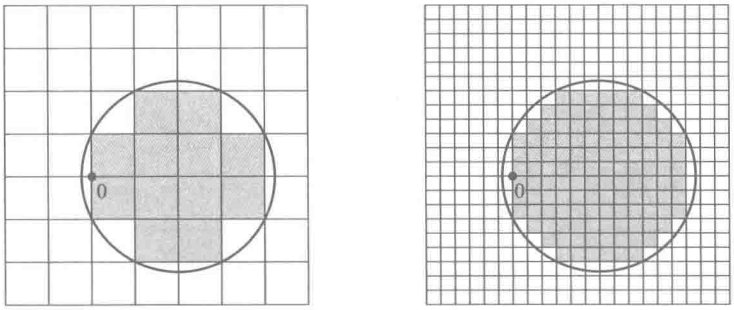
$$\begin{aligned}T(\boldsymbol{P}+S)\text{ 的面积} &= T(\boldsymbol{P})+T(S)\text{ 的面积}\\ &= T(S)\text{ 的面积}\quad(\text{平移不改变图形的面积})\\ &= |\det \boldsymbol{A}|\cdot\{S\text{ 的面积}\}\end{aligned}$$

$$= |\det \boldsymbol{A}| \cdot \{\boldsymbol{P} + S\text{ 的面积}\}.$$

综上，在$\mathbb{R}^2$中，对任意的平行四边形，(8.15)式均成立. 在$\mathbb{R}^3$中，(8.16)式的证明类似.

□

当把定理 8.5 推广到$\mathbb{R}^2$和$\mathbb{R}^3$中不是由直线或平面所围的区域时，必须面对如何定义和计算这个区域的面积或体积的问题. 这是微积分中研究的问题，也就是利用微元法的思想. 若 S 是一个有限面积的平面区域，则 R 可由其内部的小正方形方格来近似，当这些小正方形充分小时，S 的面积可由这些小正方形的面积之和无限地逼近. 见图 8.17.



(a) (b)

图 8.17

若 T 是一个矩阵变换. 由一个 2×2 的矩阵来确定，则 S 在 T 下的像由 S 中小正方形的像来近似. 若 S'是 S 中这些小正方形的并集，则 $T(S')$ 的面积就是 $|\det \boldsymbol{A}|$ 乘 S'的面积. 此时，$T(S')$ 的面积也逼近 $T(S)$ 的面积. 见图 8.18.

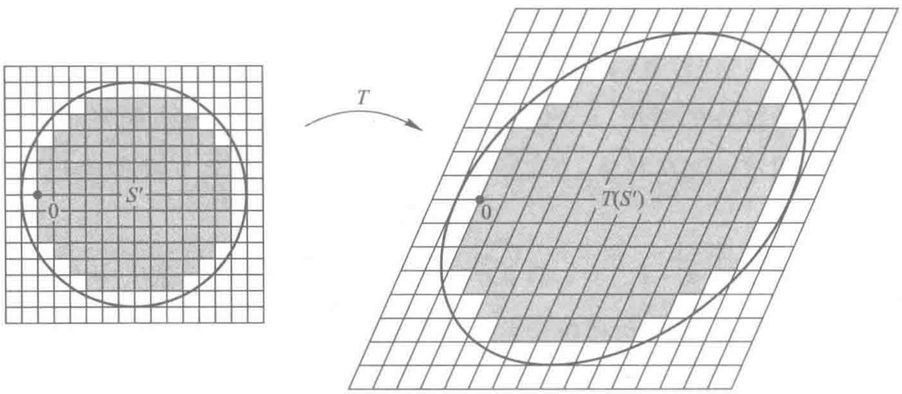


图 8.18

定理 8.6 定理 8.5 中的结论对$\mathbb{R}^2$中任意具有有限面积的区域或$\mathbb{R}^3$中具有有限体积的区域均成立.

例 8.15 若 a，b 是正数，求由方程为$\dfrac{x_1^2}{a^2}+\dfrac{x_2^2}{b^2}=1$的椭圆为边界的区域 E 的面积.

解 椭圆 E 是单位圆盘 D 在伸缩变换 T 下的像.这里 T 由矩阵 $\boldsymbol{A}=\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ 确定.因为若 $\boldsymbol{u}=\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$,$\boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$,且 $\boldsymbol{x}=\boldsymbol{A}\boldsymbol{u}$,则

$$u_1=\frac{x_1}{a},\quad u_2=\frac{x_2}{b},$$

从而得 $\boldsymbol{u}$ 在此单位圆内,即满足 $u_1^2+u_2^2\leqslant 1$.当且仅当 $\boldsymbol{x}$ 在 E 内,满足 $\left(\frac{x_1}{a}\right)^2+\left(\frac{x_2}{b}\right)^2\leqslant 1$,由定理 8.5 的推广,

$$\begin{aligned}\text{椭圆的面积} &= T(D)\text{ 的面积} \\ &= |\det \boldsymbol{A}|\cdot\{D\text{ 的面积}\} \\ &= a\cdot b\cdot \pi\cdot 1^2 \\ &= \pi ab.\end{aligned}$$

8.2.3 一般行列式的几何意义

下面,我们将行列式的几何意义推广到一般情形,即$\mathbb{R}^n$中 $m(m\leqslant n)$个向量所张成的体积.因此,首先要定义$\mathbb{R}^n$中的 m 维体积.

定义 8.3 $\mathbb{R}^n$中 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 为 m 个列向量,它们所张成的 m 维体积 $V_m(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)$按如下方式归纳地定义:

$$V_1^2(\boldsymbol{\alpha}_1)=|\boldsymbol{\alpha}_1|^2.$$

$$V_k^2(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k)=|\boldsymbol{\beta}_k|^2\cdot V_{k-1}^2(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{k-1}),$$

其中 $\boldsymbol{\beta}_k=\boldsymbol{\alpha}_k-\sum\limits_{j=1}^{k-1}\lambda_j\boldsymbol{\alpha}_j(\lambda_j\in\mathbb{R})$满足

$$\boldsymbol{\beta}_k\perp L(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{k-1}).$$

注 ① 先说明满足这个条件的 λ_j 是存在的(有可能全为 0).

事实上,由 $\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\beta}_k=0(i=1,2,\cdots,k-1)$得到 λ_j 所满足的$(k-1)$个线性方程:

$$\begin{cases}\sum\limits_{j=1}^{k-1}\lambda_j\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_j=\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_k, \\ \cdots\cdots\cdots\cdots \\ \sum\limits_{j=1}^{k-1}\lambda_j\boldsymbol{\alpha}_{k-1}^{\mathrm{T}}\boldsymbol{\alpha}_j=\boldsymbol{\alpha}_{k-1}^{\mathrm{T}}\boldsymbol{\alpha}_k.\end{cases}$$

这个线性方程组有唯一解$(\lambda_1,\cdots,\lambda_{k-1})$的充要条件是系数行列式不为零,即

$$|\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\alpha}_j|=|\boldsymbol{A}_{k-1}^{\mathrm{T}}\boldsymbol{A}_{k-1}|\neq 0,\quad i,j=1,2,\cdots,k-1.$$

$$A_{k-1}=\begin{pmatrix}a_{11} & a_{21} & \cdots & a_{k-1,1}\\ \vdots & \vdots & & \vdots\\ a_{1n} & a_{2n} & \cdots & a_{k-1,n}\end{pmatrix}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{k-1}),$$

这个矩阵的行列式不为 0,因为 $r(A_{k-1})=r(A_{k-1}^{\mathrm{T}})=k-1$,所以 $(k-1)+(k-1)-(k-1)\leqslant r(A_{k-1}^{\mathrm{T}}A_{k-1})\leqslant k-1$,因此 $r(A_{k-1}^{\mathrm{T}}A_{k-1})=k-1$,也就是满秩,于是其行列式不为 0.

② 定义 8.3 引入了向量组的"体积"概念,这与向量组的个数有关:只有一个向量时,就考虑其长度;有两个向量时,就考虑它们张成的平行四边形的面积;有三个向量时,就考虑它们张成的平行六面体的体积.以此类推.

定理 8.7 设$\mathbb{R}^n$中 m 个列向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关,而矩阵 $A=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)_{n\times m}$,则

$$V_m^2(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)=|A^{\mathrm{T}}A|.$$

证明 用数学归纳法.

当 $m=1$ 时,按定义成立;

设 $|A_{m-1}^{\mathrm{T}}A_{m-1}|=V_{m-1}^2(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{m-1})$,则

$$|A^{\mathrm{T}}A|=\begin{vmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_m\\ \vdots \qquad \vdots\\ \boldsymbol{\alpha}_m^{\mathrm{T}}\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m^{\mathrm{T}}\boldsymbol{\alpha}_m\end{vmatrix}$$

$$=\begin{vmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_1^{\mathrm{T}}\left(\boldsymbol{\beta}_m+\sum_{i=1}^{m-1}\lambda_i\boldsymbol{\alpha}_i\right)\\ \vdots & & \vdots\\ \left(\boldsymbol{\beta}_m^{\mathrm{T}}+\sum_{i=1}^{m-1}\lambda_i\boldsymbol{\alpha}_i^{\mathrm{T}}\right)\boldsymbol{\alpha}_1 & \cdots & \left(\boldsymbol{\beta}_m^{\mathrm{T}}+\sum_{i=1}^{m-1}\lambda_i\boldsymbol{\alpha}_i^{\mathrm{T}}\right)\left(\boldsymbol{\beta}_m+\sum_{i=1}^{m-1}\lambda_i\boldsymbol{\alpha}_i\right)\end{vmatrix}$$

$$=\begin{vmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_{m-1} & \sum_{i=1}^{m-1}\lambda_i\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_i\\ \vdots & & \vdots & \vdots\\ \sum_{j=1}^{m-1}\lambda_j\boldsymbol{\alpha}_j^{\mathrm{T}}\boldsymbol{\alpha}_1 & \cdots & \sum_{j=1}^{m-1}\lambda_j\boldsymbol{\alpha}_j^{\mathrm{T}}\boldsymbol{\alpha}_{m-1} & |\boldsymbol{\beta}_m|^2+\sum_{j=1}^{m-1}\lambda_j\sum_{i=1}^{m-1}\lambda_i\boldsymbol{\alpha}_j^{\mathrm{T}}\boldsymbol{\alpha}_i\end{vmatrix}.$$

把第一行乘 $-\lambda_1,\cdots$,第 $(m-1)$ 行乘 $-\lambda_{m-1}$ 后加到最后一行,然后按第 m 列展开.由归纳法假设就得到

$$|A^{\mathrm{T}}A|=|\boldsymbol{\beta}_m|^2\begin{vmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_{m-1}\\ \vdots & & \vdots\\ \boldsymbol{\alpha}_{m-1}^{\mathrm{T}}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_{m-1}^{\mathrm{T}}\boldsymbol{\alpha}_{m-1}\end{vmatrix}$$

$$= |\boldsymbol{\beta}_m|^2 V_{m-1}^2(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{m-1})$$
$$= V_m^2(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m). \qquad \square$$

特别地,在定理 8.7 中,当 $m=n$ 时,有 $V_n(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)=|\det \boldsymbol{A}|$,即行列式的绝对值等于行列式中的 n 个列(或行)向量张成的体积,这就是一般行列式的几何意义.

§ 8.3 最小二乘问题

在第二章与 § 5.5 中,我们已经讨论了线性方程组的解的理论.我们已经知道了如何判断一个方程组是否有解,当有解时,解是否唯一,解集的结构等.当非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 无解时 ,即 $r(\boldsymbol{A})\neq r(\boldsymbol{A},\boldsymbol{b})$ 时,我们称这样的方程组是**矛盾方程组**.在一些实际问题中,虽然不存在方程组的精确解,但可求得它的近似解.本节介绍如何解决此类问题.我们仍利用几何与代数相结合的方法.

8.3.1 近似解的标准——最小距离

设非齐次线性方程组为

$$\boldsymbol{A}_{m\times n}\boldsymbol{x}=\boldsymbol{b}, \tag{8.17}$$

当方程组(8.17)无解时,可考虑某种近似解.那么可采用什么样的近似标准呢？我们从几何上来考虑.设 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)$,其中 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\in\mathbb{R}^m$,则

$$\boldsymbol{Ax}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n,$$

当 $x_1,\cdots,x_n$ 跑遍全体$\mathbb{R}$ 时,$\boldsymbol{x}$ 跑遍$\mathbb{R}^n$.于是上式就生成一个子空间

$$W=L(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)=\{\boldsymbol{Ax}\mid \boldsymbol{x}\in\mathbb{R}^n\}=\{x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n\mid x_1,\cdots,x_n\in\mathbb{R}\}.$$

由于 W 是由 $\boldsymbol{A}$ 的全体列向量生成的子空间,称为 $\boldsymbol{A}$ 的**列空间**,记为 Col $\boldsymbol{A}$.因而有 $\boldsymbol{b}\in$ Col $\boldsymbol{A}$ $\Leftrightarrow\boldsymbol{Ax}=\boldsymbol{b}$ 有解.换言之,方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 无解的原因是 $\boldsymbol{b}\notin$ Col $\boldsymbol{A}$.

如图 8.19,把 Col $\boldsymbol{A}$ 想象成一个平面 W.$\boldsymbol{b}$ 虽然不在 Col $\boldsymbol{A}$ 中,但我们可以在 Col $\boldsymbol{A}$ 中找一个与 $\boldsymbol{b}$ 最接近的向量来近似它.这就是 $\boldsymbol{b}$ 在 Col $\boldsymbol{A}=W$ 上的正交投影向量

$$\boldsymbol{\beta}=(\boldsymbol{b})_W.$$

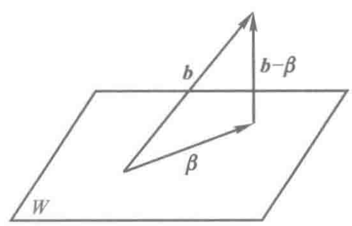


图 8.19

因 $\boldsymbol{\beta}\in W=$ Col $\boldsymbol{A}$,设 $\boldsymbol{\beta}=\boldsymbol{Ax}_0$,则 $\boldsymbol{x}_0$ 就是所要的近似解,即

$$|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_0|=\min_{\boldsymbol{x}\in\mathbb{R}^n}|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}|,$$

两边平方得

$$|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_0|^2=\min_{\boldsymbol{x}\in\mathbb{R}^n}|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}|^2.$$

再表示为内积的形式为

$$(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_0,\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_0)=\min_{\boldsymbol{x}\in\mathbb{R}^n}(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x},\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}). \tag{8.18}$$

从这个意义上讲,$\boldsymbol{x}_0$ 是方程组(8.17)的最优近似解,称为“**最小二乘解**”.所谓“最小二乘”源于 $|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}|^2$ 是平方和.

下面,我们来讨论如何求这个最优近似解.

考虑向量 $\boldsymbol{b}-\boldsymbol{\beta}$,它是与 Col $\boldsymbol{A}$ 垂直的向量,即 $(\boldsymbol{b}-\boldsymbol{\beta})\perp$ Col $\boldsymbol{A}$.故

$$(\boldsymbol{b}-\boldsymbol{\beta})\perp\boldsymbol{\alpha}_i\quad(i=1,2,\cdots,n).$$

用标准内积表示为

$$(\boldsymbol{b}-\boldsymbol{\beta},\boldsymbol{\alpha}_i)=\boldsymbol{\alpha}_i^{\mathrm{T}}(\boldsymbol{b}-\boldsymbol{\beta})=0,\quad i=1,2,\cdots,n.$$

将上述 n 个式子按行排列得

$$\begin{pmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{\alpha}_n^{\mathrm{T}}\end{pmatrix}(\boldsymbol{b}-\boldsymbol{\beta})=\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{b}-\boldsymbol{\beta})=\boldsymbol{0},$$

展开后,得

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_0=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b},$$

即近似解 $\boldsymbol{x}_0$ 是方程组

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \tag{8.19}$$

的解.方程组(8.19)通常称为“**法方程**”.

把上面结果写为定理:

定理 8.8 非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$,当 $r(\boldsymbol{A})\neq r(\boldsymbol{A},\boldsymbol{b})$ 时,

(1) 可用法方程 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 的解作为 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的最优近似解.

(2) 法方程(8.19)必有解,且当 $r(\boldsymbol{A})=n$($\boldsymbol{A}$ 列满秩)时,$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 可逆.法方程有唯一解

$$\boldsymbol{x}_0=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}. \tag{8.20}$$

证明 (1) 之前的推导已说明原方程的最优近似解是法方程的解.下面说明法方程的解必为原方程的最优近似解.设 $\boldsymbol{y}$ 满足法方程,即 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{y}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$.从而有

$$\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{y})=\boldsymbol{0}.$$

说明 $\boldsymbol{b}-\boldsymbol{A}\boldsymbol{y}$ 与 $\boldsymbol{A}$ 的列向量均正交,即 $(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{y})\perp$ Col $\boldsymbol{A}$.

由正交投影的唯一性知,$\boldsymbol{A}\boldsymbol{y}=(\boldsymbol{b})_{\mathrm{Col}\,\boldsymbol{A}}$.从而 $\boldsymbol{y}$ 是原方程组的最小二乘解.

正交投影唯一性的证明:对 $\boldsymbol{b}\notin W$,若 $z_1,z_2\in W$ 均为 $\boldsymbol{b}$ 在 W 上的正交投影,则 $\boldsymbol{y}_1=\boldsymbol{b}-z_1$,

$\boldsymbol{y}_2=\boldsymbol{b}-\boldsymbol{z}_2$ 均垂直于 W.由于 $\boldsymbol{b}=\boldsymbol{y}_1+\boldsymbol{z}_1=\boldsymbol{y}_2+\boldsymbol{z}_2$,有

$$\boldsymbol{z}_1-\boldsymbol{z}_2=\boldsymbol{y}_2-\boldsymbol{y}_1=\boldsymbol{v}.$$

上式说明 $\boldsymbol{v}\in W$ 且 $\boldsymbol{v}\perp W$,只可能 $\boldsymbol{v}=\boldsymbol{0}$,即 $\boldsymbol{z}_1=\boldsymbol{z}_2$.

(2) 要证法方程必有解,需证

$$r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}).$$

一方面,由秩的性质有

$$r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})=r(\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A},\boldsymbol{b}))\leqslant r(\boldsymbol{A}^{\mathrm{T}})=r(\boldsymbol{A});$$

另一方面,

$$r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})\geqslant r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}).$$

由例 5.14 知,$r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=r(\boldsymbol{A})$,因而

$$r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})\geqslant r(\boldsymbol{A}).$$

综上,$r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})$.故法方程(8.19)必有解.

又因 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 为 $n\times n$ 方阵.所以,$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 可逆$\Leftrightarrow r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=r(\boldsymbol{A})=n\Leftrightarrow\boldsymbol{A}$ 列满秩.此时法方程有唯一解

$$\boldsymbol{x}_0=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}.$$ □

例 8.16 给出以下线性方程组$\begin{cases}x+y+z=1,\\x+y+z=2\end{cases}$的最优近似解.

解 由于此方程组比较简单,一看就知道是一个矛盾方程组,因而有

$$\boldsymbol{A}=\begin{pmatrix}1&1&1\\1&1&1\end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix}1\\2\end{pmatrix}$$

$$\Rightarrow\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix}1&1\\1&1\\1&1\end{pmatrix}\begin{pmatrix}1&1&1\\1&1&1\end{pmatrix}=\begin{pmatrix}2&2&2\\2&2&2\\2&2&2\end{pmatrix},\quad \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}=\begin{pmatrix}3\\3\\3\end{pmatrix},$$

故法方程为 $x+y+z=\dfrac{3}{2}$.

用高斯消元法可得近似最优解为

$$\boldsymbol{x}_0=\begin{pmatrix}\frac{3}{2}\\0\\0\end{pmatrix}+k_1\begin{pmatrix}-1\\1\\0\end{pmatrix}+k_2\begin{pmatrix}-1\\0\\1\end{pmatrix},\quad k_1,k_2\in\mathbb{R}.$$

从几何上看(图 8.20),最优近似解是一个平面.在已知的两个平面中间.

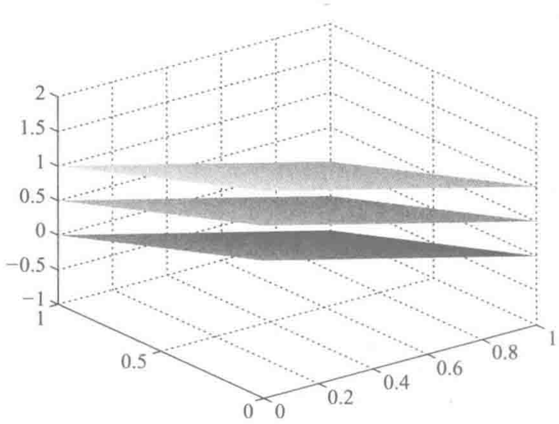


图 8.20

8.3.2　最小二乘解的应用

在一些实际问题里,人们想知道某些变量之间的函数关系,可通过实验获得数据,再从实验数据求出函数关系,这就是统计学中所谓“曲线拟合”或“回归”问题.

我们从最简单的例子出发来说明.

例 8.17　设有以下实验数据

x_i	1	2	3	4	5
y_i	1.2	1.5	2.3	2.4	3.3

求一次函数 $y=a+bx$(a,b 为待定参数),使它与实验数据的误差平方和最小.

解　假设 a,b 已确定,则当 $x=1,2,3,4,5$ 时,y 应得到如下理论值,即

$$\begin{cases} a+b=y_1^*, \\ a+2b=y_2^*, \\ a+3b=y_3^*, \\ a+4b=y_4^*, \\ a+5b=y_5^*, \end{cases}$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}=\begin{pmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \end{pmatrix},$$

记为
$$A\begin{pmatrix}a\\b\end{pmatrix}=y^*.$$

它们与实验数据的误差平方和为
$$\sum_{i=1}^{5}(y_i^*-y_i)^2,$$
表示为向量内积的形式为
$$|y^*-y|^2=\left|y-A\begin{pmatrix}a\\b\end{pmatrix}\right|^2.$$
因此求误差平方和的最小值就是求关于 $A\begin{pmatrix}a\\b\end{pmatrix}=y$ 的最小二乘解 $\begin{pmatrix}a_0\\b_0\end{pmatrix}$.由于
$$A=\begin{pmatrix}1&1\\1&2\\1&3\\1&4\\1&5\end{pmatrix},y=\begin{pmatrix}1.2\\1.5\\2.3\\2.4\\3.3\end{pmatrix},$$
则
$$A^TA=\begin{pmatrix}5&15\\15&55\end{pmatrix},\quad A^Ty=\begin{pmatrix}10.7\\37.2\end{pmatrix},\quad (A^TA)^{-1}=\frac{1}{50}\begin{pmatrix}55&-15\\-15&5\end{pmatrix},$$
得最小二乘解为
$$\begin{pmatrix}a_0\\b_0\end{pmatrix}=(A^TA)^{-1}A^Ty=\frac{1}{50}\begin{pmatrix}55&-15\\-15&5\end{pmatrix}\begin{pmatrix}10.7\\37.2\end{pmatrix}=\begin{pmatrix}0.61\\0.51\end{pmatrix}.$$
故满足条件的一次函数为
$$y=0.61+0.51x.$$

如图 8.21.

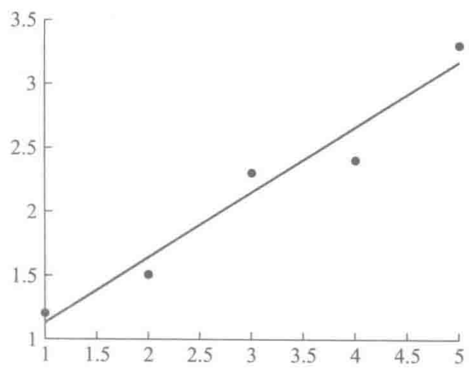


图 8.21

例 8.18　设有以下实验数据

x_i	1	2	3	4	5
y_i	6.1	10.5	18.4	26.5	38.2

求二次函数 $y=a+bx+cx^2$(a,b,c 为待定参数),使它与实验数据的误差平方和最小.

解　同上例,可说明使得误差平方和最小的二次函数系数$(a_0,b_0,c_0)^{\mathrm{T}}$,就是如下线性方程组的最小二乘解:

$$\begin{cases} a+b+c=6.1, \\ a+2b+4c=10.5, \\ a+3b+9c=18.4, \\ a+4b+16c=26.5, \\ a+5b+25c=38.2. \end{cases}$$

设

$$\boldsymbol{A}=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{pmatrix}, \quad \boldsymbol{b}=\begin{pmatrix} 6.1 \\ 10.5 \\ 18.4 \\ 26.5 \\ 38.2 \end{pmatrix},$$

则

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix} 5 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{pmatrix}, \quad \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}=\begin{pmatrix} 99.8 \\ 379.3 \\ 1\,592.7 \end{pmatrix}, \quad (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}=\begin{pmatrix} 4.6 & -3.3 & 0.5 \\ -3.3 & \frac{18.7}{7} & -\frac{3}{7} \\ 0.5 & -\frac{3}{7} & \frac{1}{14} \end{pmatrix},$$

得最小二乘解为

$$\boldsymbol{x}_0=\begin{pmatrix} a_0 \\ b_0 \\ c_0 \end{pmatrix}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}=\begin{pmatrix} 3.74 \\ 1.35 \\ 1.11 \end{pmatrix}.$$

所以误差平方和最小的二次函数为

$$y=3.74+1.35x+1.11x^2.$$

如图 8.22.

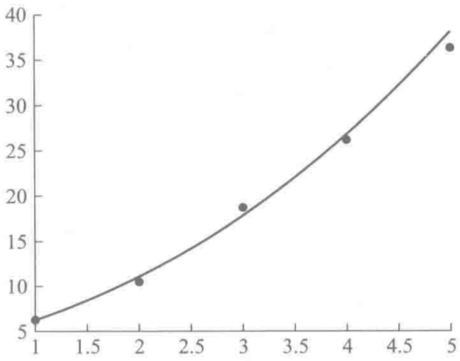


图 8.22

*8.3.3　最小二乘问题的微积分推导

除了用几何中正交投影的方法,我们也可用多元微积分中求极值的方法来讨论最小二乘的问题.设矛盾线性方程组为 $\boldsymbol{A}_{m\times n}\boldsymbol{x}=\boldsymbol{b}$,考虑平方误差函数为

$$f(x_1,\cdots,x_n)=\sum_{i=1}^{m}(a_{i1}x_1+\cdots+a_{in}x_n-b_i)^2 \tag{8.21}$$

(也可令 $f(\boldsymbol{x})=|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}|^2$).

根据多元函数极值的必要条件,平方误差函数取得最小值时对应的变量 $x_1,x_2,\cdots,x_n$ 应满足以下条件

$$\frac{\partial f}{\partial x_j}=0,\quad j=1,2,\cdots,n.$$

将(8.21)式代入得

$$2\sum_{i=1}^{m}a_{ij}(a_{i1}x_1+\cdots+a_{in}x_n-b_i)=0,\quad j=1,2,\cdots,n,$$

即

$$\sum_{i=1}^{m}a_{ij}(a_{i1}x_1+\cdots+a_{in}x_n)=\sum_{i=1}^{m}a_{ij}b_i,\quad j=1,2,\cdots,n,$$

写成矩阵与向量的形式就是法方程

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}.$$

§8.4 二次曲线与二次曲面

8.4.1 $\mathbb{R}^2$中的二次曲线及其标准方程

平面内的二次曲线又称为圆锥曲线.前一名称从代数方程而来,后一名称则源于它的几何产生方法.二次曲线在理论和实践上都是很重要的.$\mathbb{R}^2$中的二次曲线有三类:椭圆、双曲线与抛物线,上述图形也可能退化为一个点、两条相交直线、两条平行直线、两条重合直线以及各种虚轨迹.

本小节将介绍三类二次曲线的定义及标准方程.下一小节将用二次型与特征值理论讨论一般二元二次方程的化简及分类问题.

1. 椭圆的定义与方程

若平面$\mathbb{R}^2$内一个动点到两个定点的距离之和等于常量(大于两定点间的距离).那么,该动点的轨迹是什么图形呢?

设 F'与 F 为两定点,M 为动点,以 $2c$ 表示 F'与 F 间的距离(简称焦距),$2a$ 表示动点 M 到 F'与 F 的距离之和.取经过 F'与 F 的直线为 x 轴,线段 $F'F$ 的垂直平分线为 y 轴.如图 8.23 所示,在坐标系 xOy 中 $F'(-c,0)$,$F(c,0)$,$M(x,y)$,则

$$|MF'|+|MF|=\sqrt{(x+c)^2+y^2}+\sqrt{(x-c)^2+y^2}=2a,$$

化简得

$$(a^2-c^2)x^2+a^2y^2=a^2(a^2-c^2).$$

由 $a>c$,可令 $a^2-c^2=b^2(b>0)$,得

$$b^2x^2+a^2y^2=a^2b^2,$$

即

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1. \tag{8.22}$$

动点 M 的轨迹称为椭圆.F'与 F 称为椭圆的焦点.$2a$ 与 $2b$ 分别称为椭圆的长轴与短轴.(8.22)式称为椭圆的标准方程.易知,椭圆是关于原点 O、x 轴与 y 轴均对称的有界图形($|x|\leqslant a$, $|y|\leqslant b$).

2. 双曲线的定义与方程

若平面$\mathbb{R}^2$中一个动点到两个定点的距离之差的绝对值等于常量(小于两定点间的距离且不等于零),那么该动点的轨迹是什么图形?

我们设动点为 M,两定点为 F'与 F,同前建立直角坐标系,如图 8.24.

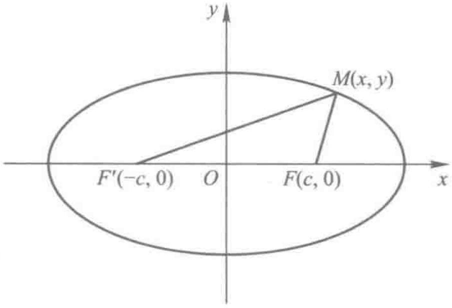


图 8.23

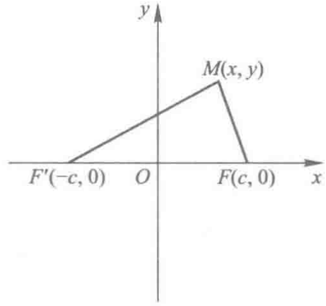


图 8.24

设 $\big||MF'|-|MF|\big|=2a$, $|F'F|=2c$,则

$$\sqrt{(x+c)^2+y^2}-\sqrt{(x-c)^2+y^2}=\pm 2a,$$

化简得

$$(c^2-a^2)x^2-a^2y^2=a^2(c^2-a^2).$$

由 $a<c$,可设 $c^2-a^2=b^2(b>0)$,得

$$b^2x^2-a^2y^2=a^2b^2,$$

即

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1. \tag{8.23}$$

动点 M 的轨迹,由方程(8.23)定义 ,称为双曲线.F',F 称为双曲线的焦点,方程(8.23)称为双曲线的标准方程.易知,双曲线是关于原点 O、x 轴、y 轴均对称的无界图形.x 轴(焦点所在的轴)称为双曲线(8.23)的实对称轴,y 轴称为双曲线(8.23)的虚对称轴.设 $A(a,0)$,$A'(-a,0)$,$B(0,b)$,$B'(0,-b)$,过 A,A',B,B'画一个中心在原点 O、边与轴平行且边长为 $2a$ 与 $2b$ 的矩形 $CDEF$(图 8.25).矩形的两条对角线 CE 与 DF 所在的两条直线,称为双曲线(8.23)的渐近线,它们满足方程$\frac{x^2}{a^2}-\frac{y^2}{b^2}=0$,这是双曲线退化的情形.

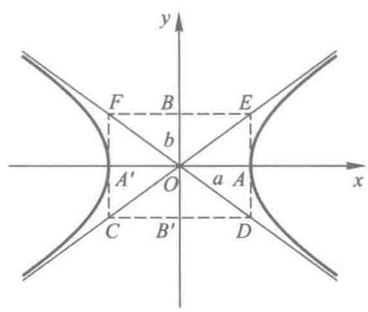


图 8.25

3. 抛物线的定义与方程

若$\mathbb{R}^2$中一动点到一定点和一条定直线的距离相等(假设定点不在定直线上),那么该动点的轨迹是什么图形?

设定点为 F,定直线为 l,动点为 M,以过点 F 且与直线 l 垂直的直线为 x 轴,x 轴与 l 相

交于 E,线段 EF 的垂直平分线为 y 轴.如图 8.26 所示.令 $|EF|=p$,则 $F=\left(\frac{p}{2},0\right)$,$l$ 的方程为 $x=-\frac{p}{2}$.设动点 $M(x,y)$,作 $MN\perp l$,垂足为 N,则 $|MF|=|MN|$,即

$$\sqrt{\left(x-\frac{p}{2}\right)^2+y^2}=\left|x+\frac{p}{2}\right|,$$

化简得

$$y^2=2px. \tag{8.24}$$

动点 M 的轨迹称为抛物线,由方程(8.24)定义.定点 F 称为抛物线(8.24)的焦点.直线 l 称为抛物线(8.24)的准线,原点 O 在抛物线(8.24)上,称为顶点,它是抛物线上距准线与焦点最近的点.易知,抛物线(8.24)仅关于 x 轴对称,无中心且无界.

4. 圆锥曲线命名由来

一个平面与一个圆锥面(具体见 8.4.3 节)相截,由于截面位置不同,可以得到各种不同的截线.特别地,椭圆、双曲线、抛物线都可由此截得.这就是圆锥曲线命名的由来.具体如下:

用一个不经过圆锥面顶点的平面截这圆锥面.设截面和锥面的轴成 θ 角,锥面的半顶角为 α(如图 8.27),则

(i) 当 $\alpha<\theta\leqslant\frac{\pi}{2}$ 时,截线是椭圆.特别当 $\theta=\frac{\pi}{2}$ 时,截线是圆.

(ii) 当 $0\leqslant\theta<\alpha$ 时,截线是双曲线.

(iii) 当 $\theta=\alpha$ 时,截线是抛物线.

(证明略.)

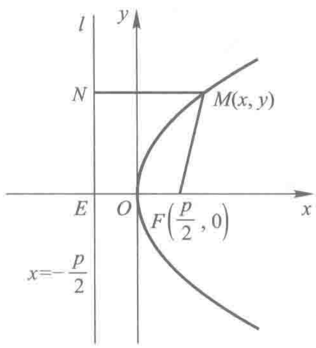


图 8.26

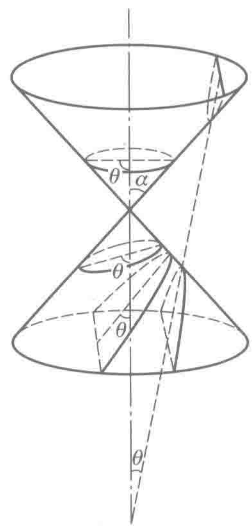


图 8.27

8.4.2 二次曲线一般方程的化简与分类

1. 一般方程的化简

在$\mathbb{R}^2$中,关于 x_1,x_2 的二元二次方程

$$a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2+2a_1x_1+2a_2x_2+a_0=0, \tag{8.25}$$

其中二次项系数不全为零,称为二次曲线的一般方程,它在$\mathbb{R}^2$中对应什么样的曲线呢?

首先,我们将(8.25)式改写成矩阵与(分块)矩阵的形式,即

$$\begin{aligned} f(x_1,x_2) &= (x_1,x_2,1)\begin{pmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a_0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = (\boldsymbol{x}^{\mathrm{T}},1)\begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 \end{pmatrix}\begin{pmatrix} \boldsymbol{x} \\ 1 \end{pmatrix} \\ &= \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}+2\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{x}+a_0=0, \end{aligned} \tag{8.26}$$

其中 $\boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$,$\boldsymbol{\alpha}=\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$,$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$,其中二次部分为

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2. \tag{8.27}$$

其次,我们将一般方程化为标准方程,方法为坐标系的平移与旋转,其中坐标系的旋转对应二次部分的正交对角化,即消去交叉项,而坐标系的平移对应配方法,消去一次项.

具体来说,设 $\boldsymbol{P}$ 为自然基经过旋转而得新基的过渡矩阵,则由前面讨论知,$\boldsymbol{P}$ 为正交矩阵,且由 $\boldsymbol{A}$ 为实对称矩阵知 $\det\boldsymbol{P}=1$.由坐标变换公式有 $\boldsymbol{P}\boldsymbol{y}=\boldsymbol{x}$,代入二次部分得

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{y}. \tag{8.28}$$

若 $\boldsymbol{P}$ 正好为定理 6.4 中的正交矩阵,则有

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix},$$

其中 λ_1,λ_2 为 $\boldsymbol{A}$ 的特征值,从而有换元后的方程

$$\begin{aligned} &\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y}+2\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{y}+a_0 \\ &=\lambda_1y_1^2+\lambda_2y_2^2+2b_1y_1+2b_2y_2+a_0=0. \end{aligned}$$

从而消去了交叉项 x_1x_2.(问题:如何保证正交阵 $\boldsymbol{P}$ 的行列式为 1?)

当 $\lambda_i\neq 0(i=1,2)$时,可用配方法消去一次项,得

$$\lambda_1\left(y_1+\frac{b_1}{\lambda_1}\right)^2+\lambda_2\left(y_2+\frac{b_2}{\lambda_2}\right)^2+c_0=\lambda_1z_1^2+\lambda_2z_2^2+c_0=0. \tag{8.29}$$

若 λ_1 与 λ_2 同号,上式可化为

$$\frac{z_1^2}{a^2}+\frac{z_2^2}{b^2}=1,0\text{ 或 }-1,$$

分别对应椭圆、点与虚椭圆(在实平面上不存在).

若 λ_1 与 λ_2 异号,上式可化为

$$\frac{z_1^2}{a^2}-\frac{z_2^2}{b^2}=\pm 1\text{ 或 }0,$$

分别对应双曲线与一对相交的直线.

当 λ_1 与 λ_2 中有一个为 0,另一个非零(不妨设 $\lambda_1\neq 0,\lambda_2=0$),则可由配方法消去 y_1 项,再将常数项吸收到 y_2 项中得

$$z_1^2+\frac{2b_1}{\lambda_1}z_2=0\quad\text{或 }z_1^2=\pm 1\text{ 或 }z_1^2=0,$$

分别对应抛物线、一对平行直线、一对虚平行直线(在实平面上不存在)或一对重合直线.

例 8.19 把二次曲线的一般方程

$$5x_1^2+4x_1x_2+2x_2^2-24x_1-12x_2+18=0 \tag{8.30}$$

化为标准方程,并作其图形.

解 记二次项系数对应的矩阵 $\boldsymbol{A}=\begin{pmatrix}5&2\\2&2\end{pmatrix}$,求得 $\boldsymbol{A}$ 的特征值为 $\lambda_1=6,\lambda_2=1$,其对应的单位特征向量

$$\boldsymbol{v}_1=\frac{1}{\sqrt{5}}\begin{pmatrix}2\\1\end{pmatrix},\quad \boldsymbol{v}_2=\frac{1}{\sqrt{5}}\begin{pmatrix}-1\\2\end{pmatrix}.$$

记 $\boldsymbol{P}=\begin{pmatrix}\frac{2}{\sqrt{5}}&-\frac{1}{\sqrt{5}}\\\frac{1}{\sqrt{5}}&\frac{2}{\sqrt{5}}\end{pmatrix}$,$\boldsymbol{X}=\boldsymbol{PY}$,即

$$\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}\frac{2}{\sqrt{5}}&-\frac{1}{\sqrt{5}}\\\frac{1}{\sqrt{5}}&\frac{2}{\sqrt{5}}\end{pmatrix}\begin{pmatrix}y_1\\y_2\end{pmatrix}, \tag{8.31}$$

则方程(8.30)中的一次项与常数项 $-24x_1-12x_2+18$ 化为

$$(-24,\ -12)\begin{pmatrix}x_1\\x_2\end{pmatrix}+18=-12\sqrt{5}y_1+18,$$

于是原方程化为 $6y_1^2+y_2^2-12\sqrt{5}y_1+18=0$,配方得

$$6(y_1-\sqrt{5})^2+y_2^2-12=0. \tag{8.32}$$

再将坐标轴平移:令

$$z_1 = y_1 - \sqrt{5}, \quad z_2 = y_2, \tag{8.33}$$

则原方程可化为标准方程

$$\frac{z_1^2}{2} + \frac{z_2^2}{12} = 1, \tag{8.34}$$

所以方程(8.30)的图形为椭圆.长轴在 z_2 轴上,长半轴为 $2\sqrt{3}$,短半轴为$\sqrt{2}$(如图 8.28).

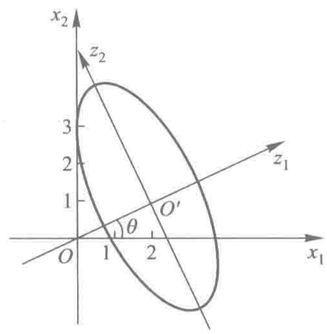


图 8.28

这里通过坐标轴旋转和平移将(8.30)化为(8.34)的坐标变换公式.由(8.31),(8.33)可得

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} z_1 + \sqrt{5} \\ z_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}}z_1 - \frac{1}{\sqrt{5}}z_2 + 2 \\ \frac{1}{\sqrt{5}}z_1 + \frac{2}{\sqrt{5}}z_2 + 1 \end{pmatrix}.$$

2. 二次曲线的不变量

从上面讨论可见,任何一条二次曲线的一般方程都可通过坐标系的旋转和平移的坐标替换,将其化为标准方程.同一条曲线在不同的直角坐标系中有不同的方程表示.这样,我们自然会提出一个问题:这些代表同一曲线的不同方程,有没有某种共性? 换言之,这种共性是不随坐标系的改变而变化的,而刻画这种共性的量,称为**二次曲线的不变量**.

定理 8.9 二次曲线的方程(8.25)式,在直角坐标替换(旋转、平移)下,有以下三个不变量:

$$I_1 = a_{11} + a_{22} = \mathrm{Tr}(\boldsymbol{A}),$$

$$I_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} = \det(\boldsymbol{A}),$$

$$I_3 = \begin{vmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a_0 \end{vmatrix} = \det\begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 \end{pmatrix}.$$

证明 (1) 先证在坐标系的旋转变换下 I_1, I_2, I_3 不变.设坐标轴旋转的变换对应坐标替换公式为

$$\boldsymbol{\alpha} = \boldsymbol{Q}\boldsymbol{y} \quad (\boldsymbol{Q}\text{ 为正交阵}),$$

从而方程变为

$$(\boldsymbol{y}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}, 1)\begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 \end{pmatrix}\begin{pmatrix} \boldsymbol{Q}\boldsymbol{y} \\ 1 \end{pmatrix} = (\boldsymbol{y}^{\mathrm{T}}, 1)\begin{pmatrix} \boldsymbol{Q}^{\mathrm{T}} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{pmatrix}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 \end{pmatrix}\begin{pmatrix} \boldsymbol{Q} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{pmatrix}\begin{pmatrix} \boldsymbol{y} \\ 1 \end{pmatrix}$$

$$= (\boldsymbol{y}^{\mathrm{T}}, 1)\begin{pmatrix} \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q} & \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{Q} & a_0 \end{pmatrix}\begin{pmatrix} \boldsymbol{y} \\ 1 \end{pmatrix}.$$

由于特征值是相似变换的不变量,于是在新变量下,

$$I_1' = \mathrm{Tr}(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}) = \lambda_1 + \lambda_2 = \mathrm{Tr}(\boldsymbol{A}) = I_1,$$

$$I_2' = \det(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}) = \lambda_1\lambda_2 = \det(\boldsymbol{A}) = I_2,$$

$$I_3' = \begin{vmatrix} \boldsymbol{Q}^{\mathrm{T}} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{vmatrix}\begin{vmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 \end{vmatrix}\begin{vmatrix} \boldsymbol{Q} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{vmatrix} = |\boldsymbol{Q}|^2 \cdot I_3 = I_3.$$

故 I_1, I_2, I_3 均不变.特别还有 I_1 为特征值之和,I_2 为特征值之积.

(2) 再证在坐标系平移变换下 I_1, I_2, I_3 不变.设平移的坐标变换公式为

$$\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{\beta} \quad (\boldsymbol{\beta} \in \mathbb{R}^2),$$

从而方程变为

$$\begin{aligned} &(\boldsymbol{y}^{\mathrm{T}} + \boldsymbol{\beta}^{\mathrm{T}}, 1)\begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 \end{pmatrix}\begin{pmatrix} \boldsymbol{y} + \boldsymbol{\beta} \\ 1 \end{pmatrix} \\ = &(\boldsymbol{y}^{\mathrm{T}}, 1)\begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 \end{pmatrix}\begin{pmatrix} \boldsymbol{y} \\ 1 \end{pmatrix} + 2(\boldsymbol{\beta}^{\mathrm{T}}, 0)\begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 \end{pmatrix}\begin{pmatrix} \boldsymbol{y} \\ 1 \end{pmatrix} + (\boldsymbol{\beta}^{\mathrm{T}}, 0)\begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 \end{pmatrix}\begin{pmatrix} \boldsymbol{\beta} \\ 0 \end{pmatrix} \\ = &(\boldsymbol{y}^{\mathrm{T}}, 1)\begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} + \boldsymbol{A}\boldsymbol{\beta} \\ \boldsymbol{\alpha}^{\mathrm{T}} + \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A} & a_0 + 2\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha} + \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta} \end{pmatrix}\begin{pmatrix} \boldsymbol{y} \\ 1 \end{pmatrix}, \end{aligned}$$

由于二次项不改变,故 I_1, I_2 不变.对于 I_3,我们可将原来的(3 阶)系数矩阵经过如下初等变换得到新的系数矩阵.由行列式的性质知 I_3 不变.

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 \end{pmatrix} \xrightarrow{\text{②列}+\text{①列}\times\boldsymbol{\beta}} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} + \boldsymbol{A}\boldsymbol{\beta} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_0 + \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta} \end{pmatrix} \xrightarrow{\text{②行}+\boldsymbol{\beta}^{\mathrm{T}}\times\text{①行}} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{\alpha} + \boldsymbol{A}\boldsymbol{\beta} \\ \boldsymbol{\alpha}^{\mathrm{T}} + \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A} & a_0 + 2\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha} + \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta} \end{pmatrix}.$$

综上,在坐标系旋转与平移的坐标替换下,二次曲线方程中的 I_1, I_2, I_3 为不变量. □

3. 利用不变量判定二次曲线的类型

下面,我们将利用 I_1, I_2, I_3 将二次曲线的最简单方程表示出来,进而可判定其对应的类型.

在上面定理的证明中,由相似矩阵特征值的不变性,我们知道

$$I_1 = \lambda_1 + \lambda_2, \quad I_2 = \lambda_1 \cdot \lambda_2.$$

故首先可用 I_2 是否为 0,将椭圆型、双曲线型(有心曲线)与抛物线型(无心曲线)分开.

对有心曲线类型,最简方程为 $\lambda_1 y_1^2 + \lambda_2 y_2^2 + b_0 = 0$.

$$I_2 > 0 \Leftrightarrow \lambda_1, \lambda_2 \text{ 同号}; I_2 < 0 \Leftrightarrow \lambda_1, \lambda_2 \text{ 异号}.$$

进而 $I_3=\begin{vmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & b_0\end{vmatrix}=I_2b_0\Rightarrow b_0=\dfrac{I_3}{I_2}$,即

$$\lambda_1y_1^2+\lambda_2y_2^2+\frac{I_3}{I_2}=0.$$

根据不变量判别二次曲线所属类型的方法如下：

(1) 当 $I_2=\lambda_1\lambda_2>0$ 时(曲线为椭圆型)；

若 I_3 与 $I_1=\lambda_1+\lambda_2$ 异号(即$\dfrac{I_3}{I_2}$与 λ_1,λ_2 异号),则曲线是椭圆；

若 I_3 与 I_1 同号,则曲线是虚椭圆；

若 $I_3=0$,则曲线是一个点.

(2) 当 $I_2=\lambda_1\lambda_2<0$ 时(曲线为双曲型)；

若 $I_3\neq 0$,则曲线是双曲线；

若 $I_3=0$,则曲线是一对相交直线.

对无心曲线类型,不妨设 $\lambda_1=0,\lambda_2\neq 0$,最简方程为 $\lambda_2y_2^2+2b_1y_1=0$,其不变量为

$$I_1=\lambda_2,\quad I_2=\begin{vmatrix}0 & 0\\ 0 & \lambda_2\end{vmatrix}=0,\quad I_3=\begin{vmatrix}0 & 0 & b_1\\ 0 & \lambda_2 & 0\\ b_1 & 0 & 0\end{vmatrix}=-\lambda_2b_1^2=-I_1b_1^2.$$

于是最简方程为

$$I_1y_2^2\pm 2\sqrt{\frac{-I_3}{I_1}}y_1=0.$$

故,若 $I_2=0,I_3\neq 0$,则曲线为抛物线;若 $I_2=0,I_3=0$,则方程可化为 $I_1y_2^2+b_0=0$(证明略).

当 b_0 与 I_1 异号时,为一对平行直线;当 b_0 与 I_1 同号时,为一对虚平行直线.当 $b_0=0$ 时,为一对重合直线.

用不变量将二次曲线分类的情况列表 8.1 如下：

例 8.20 判断下列二次曲线类型,把方程化为最简形式,并确定曲线的形状.

(1) $x_1^2+4x_1x_2+4x_2^2+12x_1-x_2+1=0$.

(2) $4x_1^2-4x_1x_2+x_2^2+2x_1-x_2-2=0$.

(3) $x_1^2-3x_1x_2+x_2^2+10x_1-10x_2+21=0$.

解 (1) $I_2=\begin{vmatrix}1 & 2\\ 2 & 4\end{vmatrix}=0$,这是抛物型曲线.此时

表 8.1

<table>
<tr><th colspan="3">不变量</th><th>曲线形状</th><th>化简后的方程</th></tr>
<tr><td rowspan="4">$I_2 \neq 0$
有心
二次
曲线</td><td rowspan="2">$I_2 > 0$
椭圆型</td><td>$I_3 \neq 0$</td><td>$I_3 I_1 < 0$ 为椭圆
$I_3 I_1 > 0$ 为虚椭圆</td><td rowspan="4">$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \frac{I_3}{I_2} = 0$</td></tr>
<tr><td>$I_3 = 0$</td><td>一个点</td></tr>
<tr><td rowspan="2">$I_2 < 0$
双曲型</td><td>$I_3 \neq 0$</td><td>双曲线</td></tr>
<tr><td>$I_3 = 0$</td><td>一对相交直线</td></tr>
<tr><td rowspan="2">$I_2 = 0$
无心
二次
曲线</td><td rowspan="2">$I_1 \neq 0$
抛物型</td><td>$I_3 \neq 0$</td><td>抛物线</td><td>$I_1 y_2^2 + 2\sqrt{-\frac{I_3}{I_1}}\, y_1 = 0$</td></tr>
<tr><td>$I_3 = 0$</td><td>$k_1 < 0$,一对平行直线*
$k_1 = 0$,一对重合直线
$k_1 > 0$,一对虚平行线</td><td>$I_1 y_2^2 + \frac{k_1}{I_1} = 0$</td></tr>
</table>

* 其中 $k_1 = \begin{vmatrix} a_{11} & a_1 \\ a_1 & a_0 \end{vmatrix} + \begin{vmatrix} a_{22} & a_2 \\ a_2 & a_0 \end{vmatrix}$ 称为二次曲线的半不变量,它在坐标轴旋转的坐标变换下不变,对于 $I_2 = I_3 = 0$ 的二次曲线,在坐标轴平移的坐标变换下也不变.

$$I_1 = 1 + 4 = 5, \quad I_3 = \begin{vmatrix} 1 & 2 & 6 \\ 2 & 4 & -\frac{1}{2} \\ 6 & -\frac{1}{2} & 1 \end{vmatrix} = -\frac{625}{4},$$

所以这是一条抛物线,化简后的方程为

$$5y_2^2 + 2y_1\sqrt{\frac{625/4}{5}} = 0 \quad 或 \quad 5y_2^2 - 2y_1\sqrt{\frac{625/4}{5}} = 0,$$

即 $y_2^2 + \sqrt{5}\, y_1 = 0$ 或 $y_2^2 - \sqrt{5}\, y_1 = 0$.

因此这条抛物线的焦参数 $p = \frac{\sqrt{5}}{2}$.

(2) $I_2 = \begin{vmatrix} 4 & -2 \\ -2 & 1 \end{vmatrix} = 0$ 这是抛物型曲线.此时

$$I_1 = 1 + 4 = 5, \quad I_3 = \begin{vmatrix} 4 & -2 & 1 \\ -2 & 1 & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -2 \end{vmatrix} = 0,$$

$$k_1=\begin{vmatrix}4 & 1\\ 1 & -2\end{vmatrix}+\begin{vmatrix}1 & -\frac{1}{2}\\ -\frac{1}{2} & -2\end{vmatrix}=-\frac{45}{4}<0,$$

所以这是两条平行直线,化简后的方程为

$$5y_2^2-\frac{9}{4}=0 \quad 即 \quad y=\pm\frac{3\sqrt{5}}{10}.$$

(3) $I_2=\begin{vmatrix}1 & -\frac{3}{2}\\ -\frac{3}{2} & 1\end{vmatrix}=-\frac{5}{4}<0$,这是双曲型曲线.此时

$$I_3=\begin{vmatrix}1 & -\frac{3}{2} & 5\\ -\frac{3}{2} & 1 & -5\\ 5 & -5 & 21\end{vmatrix}=-\frac{5}{4}, \quad I_1=1+1=2.$$

因为 $I_3\neq 0$,所以这是双曲线,解特征方程

$$|\lambda \boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1 & \frac{3}{2}\\ \frac{3}{2} & \lambda-1\end{vmatrix}=(\lambda-1)^2-\frac{9}{4}=\left(\lambda+\frac{1}{2}\right)\left(\lambda-\frac{5}{2}\right)=0,$$

得特征根 $\lambda_1=-\frac{1}{2},\lambda_2=\frac{5}{2}$.

又$\frac{I_3}{I_2}=1$,于是化简后的方程为

$$-\frac{1}{2}y_1^2+\frac{5}{2}y_2^2+1=0 \text{ 即} \frac{y_1^2}{2}-\frac{y_2^2}{\frac{2}{5}}=1.$$

所以这条双曲线的实半轴 $a=\sqrt{2}$,虚半轴 $b=\sqrt{\frac{2}{5}}=\frac{\sqrt{10}}{5}$.

8.4.3 3维空间中的二次曲面

我们知道,在$\mathbb{R}^2$中一个二元二次方程代表平面上一条二次曲面,分为三类(见表 8.2):

表 8.2 非退化非虚的二元二次方程的几何分类

类型	标准方程
椭圆	$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$
抛物线	$y^2=px$
双曲线	$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$

任何一个二元二次方程 $ax^2+bxy+cy^2+dx+ty+f=0$ 均可通过坐标平移和旋转化为标准方程.

那么,在$\mathbb{R}^3$中,一个三元二次方程对应什么样的图形呢?本节来讨论这个问题.

1. 典型的二次曲面

(1) 柱面

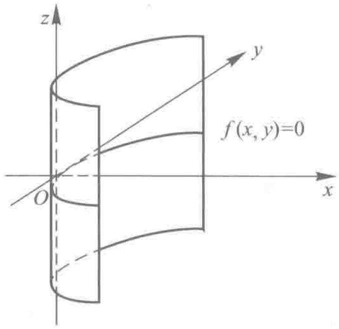
一个三元方程 $f(x,y,z)=0$ 中如果缺一变量,例如

$$f(x,y)=0, \tag{8.35}$$

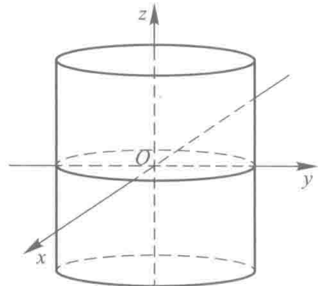
在$\mathbb{R}^3$中表明 z 坐标不受约束,因此我们可以在 xy 平面上画出曲线 $f(x,y)=0$,则过此曲线平行于 z 轴的直线上任意点均满足方程(8.35),所以方程(8.35)在$\mathbb{R}^3$中的图形是一个柱面,xy 平面上的曲线 $f(x,y)=0$ 称为此柱面的准线,平行于 z 轴的直线称为此柱面的母线(图 8.29(a)).

类似地,$f(x,z)=0$ 在$\mathbb{R}^3$中的图形是一个柱面,它的准线是 xz 平面上的曲线 $f(x,z)=0$,它的母线是平行于 y 轴的直线.$f(y,z)=0$ 在$\mathbb{R}^3$中的图形是一个柱面,它的准线是 yz 平面上的曲线 $f(y,z)=0$,它的母线是平行于 x 轴的直线.

同一个方程 $f(x,y)=0$,在$\mathbb{R}^2$中,它只是一条曲线.但在$\mathbb{R}^3$中的图形是一个柱面,例如 $x^2+y^2=r^2$,在 xy 平面上是一个圆,而在$\mathbb{R}^3$中的图形是一个圆柱面(图 8.29(b)).



(a)柱面



(b)圆柱面

图 8.29 一般柱面与圆柱面

(2) 锥面

给定曲面方程为

$$x^2+y^2=z^2, \tag{8.36}$$

确定它的图形.

现在介绍一种确定曲面的方法——平面截割法,先用平面 $z=c$ 与曲面(8.36)相交得曲线

$$\begin{cases} x^2+y^2=z^2, \\ z=c, \end{cases}$$

即

$$\begin{cases} x^2+y^2=c^2, \\ z=c, \end{cases}$$

取不同的 c,得到一系列不同半径的圆.为了看清半径的变化,我们再用平面 $x=0$ 与曲面(8.36)截割,得到

$$\begin{cases} y^2=z^2, \\ x=0 \end{cases} \quad 即 \quad \begin{cases} y=\pm z, \\ x=0 \end{cases}$$

是两条对称的直线.

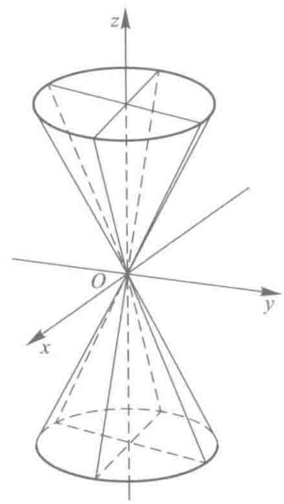


图 8.30 圆锥面

所以方程(8.36)的图形是一个锥面(图 8.30).

(3) 椭球面

椭球面的标准方程是$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$.用平面截割法,用 $x=0,y=0,z=0$ 三个坐标平面去截割,得到三个椭圆,是椭球的三个大圆.很容易画出它的图形(图 8.31).

(4) 椭圆抛物面

椭圆抛物面的标准方程是$\frac{x^2}{a^2}+\frac{y^2}{b^2}=z$.用平面截割法,用 $x=0,y=0,z=c$ 三个平面去截割,很容易得到它的图形(图 8.32).

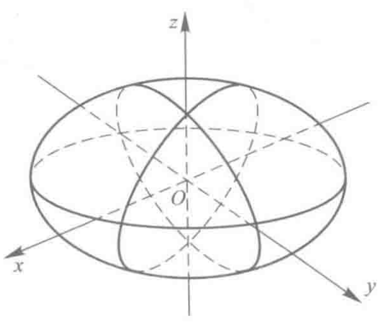


图 8.31 椭球面

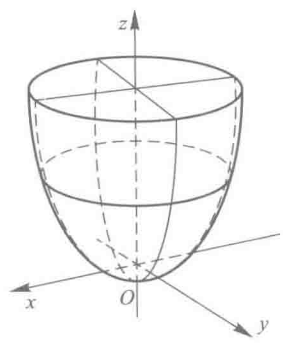


图 8.32 椭圆抛物面

(5) 单叶双曲面

单叶双曲面的标准方程是$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$.用平面截割法,用 $x=r,y=s,z=t$ 三个平面去截割,很容易得到它的图形(图 8.33).

(6) 双叶双曲面

双叶双曲面的标准方程是$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$.用平面截割法,用 $x=r,y=s,z=t$ 三个平面去截割,很容易得到它的图形(图 8.34).

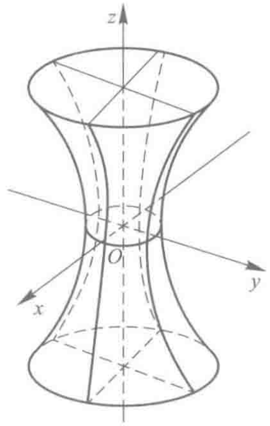


图 8.33 单叶双曲面

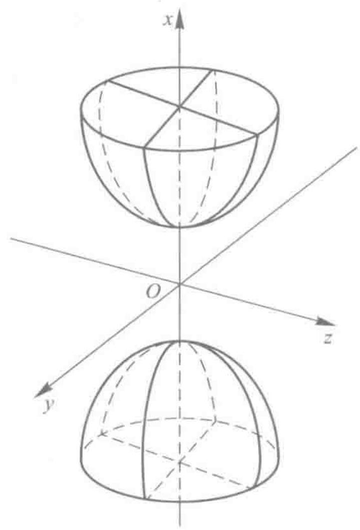


图 8.34 双叶双曲面

注 锥面实际上是单叶或双叶双曲面的渐近平面,可视为退化情形.

(7) 双曲抛物面(马鞍面)

双曲抛物面的标准方程是$\frac{x^2}{a^2}-\frac{y^2}{b^2}=cz,c\neq 0$.用平面截割法,用 $x=r,y=s,z=t$ 三个平面去截割,设 $c>0$,取平面 $z=t_1>0$ 去截割,得到实轴平行于 x 轴的双曲线,取平面 $z=t_2<0$ 去截割,得到实轴平行于 y 轴的双曲线;用平面 $y=0$ 去截割,得到对称轴为 z 轴的开口向上的抛物线;用平面 $x=0$ 去截割,得到对称轴为 z 轴的开口向下的抛物线.由此可看出它的图形是一个马鞍形的曲面(图 8.35).

2. 生成二次曲面的典型方法

除了柱面之外,还有一些方法可由 2 维图形(直线或曲线)生成二次曲面.

(1) 旋转面

设在 xy 平面上的曲线 $f(x,y)=0$ 绕 x 轴转 360°得到一个旋转面,求此旋转面的方程.我们从曲面方程的基本概念出发,于旋转面上任取一点 P,设其坐标为 $P(x,y,z)$,则 x,y,z 所满足的关系式就是该曲面的方程.也可把点 P 看作一个动点,曲面是动点运动的轨迹.我们

用平面截割法来推导旋转面的方程,过 P 作与 yz 坐标平面平行的平面,设此平面与平面 xy 上的曲线 $f(x,y)=0$ 的交点为 $A(x_0,y_0,0)$,则由图 8.36 可得

$$y^2+z^2=y_0^2. \tag{8.37}$$

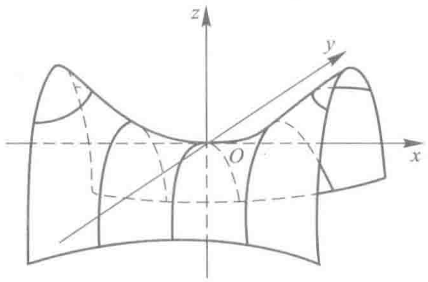


图 8.35 双曲抛物面(马鞍面)

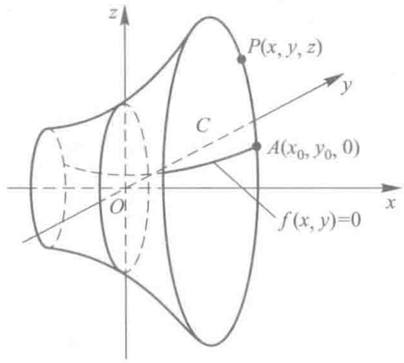


图 8.36 旋转面

由于点 $A(x_0,y_0,0)$ 在曲线上,满足 $f(x,y_0)=0$,把式(8.37)代入得

$$f(x,\ \pm\sqrt{y^2+z^2})=0,$$

这就是所要的旋转面的方程.

旋转面有广泛的应用,例如天线、探照灯的反射镜都是抛物线旋转面.

(2) 直纹面

有一类曲面,可由直线的移动来产生,这类曲面称为直纹面.例如,柱面与锥面都是直纹面.有一些曲面看起来与直线无联系,例如双曲抛物面和单叶双曲面,其实它们都是直纹面,下面来说明.

例 8.21 证明双曲抛物面和单叶双曲面是直纹面.

证明 (1) 证明双曲抛物面是直纹面.

设双曲抛物面方程为

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=cz,\quad c\neq 0. \tag{8.38}$$

将(8.38)写为 $\left(\frac{x}{a}+\frac{y}{b}\right)\left(\frac{x}{a}-\frac{y}{b}\right)=cz$,令 $\frac{x}{a}+\frac{y}{b}=\lambda$,则得 $\lambda\left(\frac{x}{a}-\frac{y}{b}\right)=cz$,设直线方程为

$$l_\lambda:\begin{cases}\dfrac{x}{a}+\dfrac{y}{b}=\lambda,\\[2ex] \lambda\left(\dfrac{x}{a}-\dfrac{y}{b}\right)=cz,\end{cases} \tag{8.39}$$

其中 λ 可取任意数.当 λ 取不同数时,由式(8.39)得到一族直线.

当 λ 连续变动时,由 l_λ 生成一个曲面,下面我们证明由式(8.39)生成的曲面就是双曲

抛物面(8.38).

首先证明双曲抛物面(8.38)上任何一点(x_0,y_0,z_0)都在直线族中的某一条直线上.令$\dfrac{x_0}{a}+\dfrac{y_0}{b}=\lambda_0$,则由式(8.38)得到$\lambda_0\left(\dfrac{x_0}{a}-\dfrac{y_0}{b}\right)=cz_0$,所以点$(x_0,y_0,z_0)$在直线$l_{\lambda_0}$上.

反之,可见直线族(8.39)中任何一条直线

$$l_{\lambda_0}:\begin{cases}\dfrac{x}{a}+\dfrac{y}{b}=\lambda_0,\\ \lambda_0\left(\dfrac{x}{a}-\dfrac{y}{b}\right)=cz\end{cases}$$

也满足方程(8.38),所以直线l_{λ_0}在曲面(8.38)上,由于λ_0可任取,故直线族(8.39)都在曲面(8.38)上.

综上,曲面(8.38)是直纹面.注:也可令$\dfrac{x}{a}-\dfrac{y}{b}=\mu$,得到另一族直线.

(2) 用类似的方法可证明单叶双曲面是直纹面. □

直纹面在实际中很有用,可用直线形的构件来构造一个曲面.由于直线是无穷长的,所以直纹面一定是无穷面.因此椭球面就不是直纹面.

3. $\mathbb{R}^3$中二次曲面的分类初步

在$\mathbb{R}^3$中一般的二次方程为

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz+b_1x+b_2y+b_3z+c=0. \tag{8.40}$$

曲面的类型主要由二次部分决定,因此可把二次部分单独拿出来讨论.

令$\boldsymbol{X}=(x,y,z)^{\mathrm{T}}$,且

$$f(\boldsymbol{X})=a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz, \tag{8.41}$$

$$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{12}&a_{22}&a_{23}\\a_{13}&a_{23}&a_{33}\end{pmatrix},$$

$$f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}.$$

其中$\boldsymbol{A}$是一个3阶实对称矩阵.由上节讨论知,存在对角矩阵$\boldsymbol{\Lambda}$和正交矩阵$\boldsymbol{P}$,使得

$$\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{\mathrm{T}}.$$

代入上式得$f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X}=(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X})^{\mathrm{T}}\boldsymbol{\Lambda}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X})$,由于$(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X},\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X})=(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X})^{\mathrm{T}}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{X}=(\boldsymbol{X},\boldsymbol{X})$.我们知道,左乘正交矩阵$\boldsymbol{P}^{\mathrm{T}}$不改变内积,从而不改变几何对象的度量(长度,夹角),即相当于旋转($|\boldsymbol{P}|=1$)与反射($|\boldsymbol{P}|=-1$).

下面对曲面作具体的分类.

设用正交变量替换$\boldsymbol{X}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{Y}$,将方程(8.41)化为标准形,即将$\boldsymbol{A}$对角化得

$$f(\boldsymbol{Y})=\lambda_1x^2+\lambda_2y^2+\lambda_3z^2,$$

其中 $\Lambda=\operatorname{diag}(\lambda_1,\lambda_2,\lambda_3)$， $\boldsymbol{Y}=(x,y,z)^{\mathrm{T}}\in\mathbb{R}^3$.

（1）若 $\lambda_1,\lambda_2,\lambda_3\neq0$ 且同号，则配方消去一次项后，总可化为

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1,0\text{ 或 }-1,$$

分别对应椭球、点或虚椭球（在实空间中不存在）.

（2）若 $\lambda_1,\lambda_2,\lambda_3\neq0$ 且异号，则消去一次项后，总可化为

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1\quad\text{或}\quad\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1\quad\text{或}\quad x^2+y^2-z^2=0,$$

分别对应单叶双曲面、双叶双曲面或锥面.

（3）若 $\lambda_1,\lambda_2,\lambda_3$ 中恰有一个为 0，另两个同号，则总可化为

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=cz\quad\text{或}\quad\frac{x^2}{a^2}+\frac{y^2}{b^2}=\pm1\quad\text{或}\quad\frac{x^2}{a^2}+\frac{y^2}{b^2}=0,$$

分别对应椭圆抛物面、椭圆柱面、虚椭圆柱面或一条直线.

（4）若 $\lambda_1,\lambda_2,\lambda_3$ 中恰有一个为 0，另两个异号，则总可化为

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=cz\quad\text{或}\quad\frac{x^2}{a^2}-\frac{y^2}{b^2}=1\quad\text{或}\quad\frac{x^2}{a^2}-\frac{y^2}{b^2}=0,$$

分别对应马鞍面、双曲线柱面或一对交叉的平面.

（5）若 $\lambda_1,\lambda_2,\lambda_3$ 中恰有两个为 0，则适当变换后总可化为

$$ax^2+cy=0\quad\text{或}\quad x^2=\pm1\quad\text{或}\quad x^2=0,$$

分别对应抛物线柱面、一对平行平面、一对虚平行平面或一对重合平面.

这样初步解决了几何空间中二次方程的分类问题.如果只要求分类，则只需求出 $\boldsymbol{A}$ 的特征值 $\lambda_1,\lambda_2,\lambda_3$ 即可判断.如要得到具体图形，则需求出具体的正交线性替换.下面举一个例子.

例 8.22 设二次方程为

$$x^2+2y^2+2z^2-4yz-2x+2\sqrt{2}y-6\sqrt{2}z+5=0,\tag{8.42}$$

准确画出它的图形.

解 先把二次部分写成如下形式

$$f(\boldsymbol{\alpha})=x^2+2y^2+2z^2-4yz,$$

因为要准确画出它的图形，必须用正交线性替换将它化为标准形，写出对应的矩阵

$$\boldsymbol{A}=\begin{pmatrix}1&0&0\\0&2&-2\\0&-2&2\end{pmatrix},$$

求出 $\boldsymbol{A}$ 的特征值为 $\lambda_1=1,\lambda_2=4,\lambda_3=0$,求出对应的特征向量为

$$\boldsymbol{\alpha}_1=(1,0,0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(0,1,-1)^{\mathrm{T}},\quad \boldsymbol{\alpha}_3=(0,1,1)^{\mathrm{T}}.$$

将它们单位化得

$$\boldsymbol{\varepsilon}_1=(1,0,0)^{\mathrm{T}},\quad \boldsymbol{\varepsilon}_2=\left(0,\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)^{\mathrm{T}},\quad \boldsymbol{\varepsilon}_3=\left(0,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)^{\mathrm{T}}.$$

令

$$\boldsymbol{Q}=\begin{pmatrix}1&0&0\\0&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\0&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix}.$$

现把方程(8.42)用矩阵表示为

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}+\boldsymbol{B}\boldsymbol{X}+5=0.$$

其中 $\boldsymbol{B}=(-2,2\sqrt{2},-6\sqrt{2})$.作正交线性替换 $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Y}$,其中 $\boldsymbol{Y}=(x_1,y_1,z_1)^{\mathrm{T}}$,得

$$\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{Y}+\boldsymbol{B}\boldsymbol{Q}\boldsymbol{Y}+5=0.$$

即 $x_1^2+4y_1^2-2x_1+8y_1-4z_1+5=0$,

进行配方,得

$$(x_1-1)^2+4(y_1+1)^2=4z_1.$$

再作可逆线性替换 $x_2=x_1-1,y_2=y_1+1,z_2=z_1$(其几何意义是坐标平移),则得

$$\frac{x_2^2}{4}+y_2^2=z_2,$$

所以二次曲面(8.42)是一个椭圆抛物面.我们可以在坐标系 $O_2-x_2y_2z_2$ 中把图形准确画出.由于从坐标系 $O-xyz$ 到坐标系 $O_2-x_2y_2z_2$,只进行了旋转和平移,因此图形除了位置有所变化外,几何形状没有任何变化.

如前讨论,$\mathbb{R}^3$ 中二次曲面一共可分为 17 类,包括复空间中的虚图形,进一步的分析本书不作深入讨论,有兴趣的读者可参阅:居余马,李海中.代数与几何.2 版.北京:高等教育出版社,2003.

习 题 八

1. 证明:$\forall \boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$,有 $(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y})\boldsymbol{y}=(\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}})\boldsymbol{x}$.

2. 求下列矩阵变换的变换矩阵:

(1) $\mathbb{R}^2$ 中关于直线 $3x-4y=0$ 对称的矩阵变换;

(2) $\mathbb{R}^2$中关于直线 $3x+4y=0$ 正交投影的矩阵变换;

(3) $\mathbb{R}^2$中绕原点逆时针旋转 $60°$的矩阵变换;

(4) $\mathbb{R}^3$中关于平面 $x+2y-2z=0$ 正交投影的矩阵变换.

3. $\mathbb{R}^n$为 n 维欧氏空间,V 是$\mathbb{R}^n$ 的一个 r 维子空间,V 有一组标准正交基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$(均为列向量),试推导关于 V 的正交投影变换矩阵和对称变换矩阵,并且以此重新推导$\mathbb{R}^2$中的关于原点的对称变换、关于过原点直线的投影变换和对称变换,以及$\mathbb{R}^3$中关于原点的对称变换、关于过原点的直线的投影变换和对称变换、关于过原点的平面的投影变换和对称变换的变换矩阵公式.

4. $\mathbb{R}^2$有基 $\boldsymbol{\alpha}_1=\begin{pmatrix}1\\1\end{pmatrix},\boldsymbol{\alpha}_2=\begin{pmatrix}0\\1\end{pmatrix}$和基 $\boldsymbol{\beta}_1=\begin{pmatrix}3\\7\end{pmatrix},\boldsymbol{\beta}_2=\begin{pmatrix}2\\5\end{pmatrix}$,线性变换 φ 在基$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)$下的矩阵为 $\begin{pmatrix}2&-1\\3&1\end{pmatrix}$,求 φ 在基$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)$下的矩阵.

5. 设 $\boldsymbol{A}=\begin{pmatrix}0&1\\-2&3\end{pmatrix},\boldsymbol{b}=\begin{pmatrix}3\\4\end{pmatrix}$,按照下面的步骤计算 $\boldsymbol{A}^{10}\boldsymbol{b}$.

(1) 设 $\boldsymbol{A}$ 的两个特征值为 u_1,u_2,其中 $u_1\leqslant u_2$,属于 u_1 的特征向量 $\boldsymbol{x}_1=\begin{pmatrix}1\\k_1\end{pmatrix}$,属于 u_2 的特征向量 $\boldsymbol{x}_2=\begin{pmatrix}1\\k_2\end{pmatrix}$,求 u_1,u_2,k_1,k_2.

(2) 将 $\boldsymbol{b}$ 在 $\boldsymbol{A}$ 的特征向量上进行分解,即求满足 $\boldsymbol{b}=s_1\boldsymbol{x}_1+s_2\boldsymbol{x}_2$ 的 s_1,s_2.

(3) 利用性质 $\boldsymbol{A}\boldsymbol{b}=\boldsymbol{A}(s_1\boldsymbol{x}_1+s_2\boldsymbol{x}_2)=s_1u_1\boldsymbol{x}_1+s_2u_2\boldsymbol{x}_2$,求 $\boldsymbol{A}^{10}\boldsymbol{b}$.

6. 判断正误:

(1) φ 为正交变换当且仅当 φ 将任意两个正交向量映成正交向量.

(2) φ 为正交变换当且仅当 φ 保持任意两个向量的夹角.

(3) φ 为正交变换当且仅当 φ 保持任意两个向量的内积.

(4) φ 为正交变换当且仅当 φ 保持任意一个向量的长度.

(5) φ 为正交变换当且仅当 φ 将任一组标准正交基映成标准正交基.

(6) φ 为正交变换当且仅当 φ 在任一组标准正交基下的矩阵为正交矩阵.

(7) φ 为正交变换当且仅当 φ 将某一组标准正交基映成标准正交基.

(8) φ 为正交变换当且仅当 φ 在某一组标准正交基下的矩阵为正交矩阵.

(9) 正交变换的合成仍然是正交变换.

(10) 正交变换的逆仍然是正交变换.

(11) 正交变换的特征值一定是 1 或 -1.

(12) 第一类正交变换一定有特征值 1.

(13) 第一类正交变换一定没有特征值 -1.

(14) $\mathbb{R}^2,\mathbb{R}^3$中第二类正交变换一定有特征值 -1.

(15) 第二类正交变换一定没有特征值 1.

7. 用行列式的几何意义解释:行列式为 0 当且仅当其行(列)向量线性相关.

8. (1) 证明:欧氏空间$\mathbb{R}^2$中的 3 个点(x_1,y_1),(x_2,y_2),(x_3,y_3)围成的三角形的面积为

$$\frac{1}{2}\left|\det\begin{pmatrix}1 & x_1 & y_1\\ 1 & x_2 & y_2\\ 1 & x_3 & y_3\end{pmatrix}\right|.$$

(2) 证明:欧氏空间$\mathbb{R}^3$中的 4 个点(x_1,y_1,z_1),(x_2,y_2,z_2),(x_3,y_3,z_3),(x_4,y_4,z_4)围成的四面体的体积为

$$\frac{1}{6}\left|\det\begin{pmatrix}1 & x_1 & y_1 & z_1\\ 1 & x_2 & y_2 & z_2\\ 1 & x_3 & y_3 & z_3\\ 1 & x_4 & y_4 & z_4\end{pmatrix}\right|.$$

(3) 试将上述结论推广到一般的 n 维欧氏空间当中.

9. 设 V 是定义了内积的线性空间,W 是 V 的子空间,$\boldsymbol{v}\in V$. 证明:如果 $\boldsymbol{v}\in W$,$\boldsymbol{v}\perp W$,则有 $\boldsymbol{v}=\mathbf{0}$.

10. 判断下列矩阵方程是否有解

$$\begin{pmatrix}a^2 & ab & ac\\ ab & b^2 & bc\\ ac & bc & c^2\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}-a\\ -b\\ -c\end{pmatrix}.$$

11. 下表提供了某厂节能降耗技术改造后生产甲产品过程中记录的产量 X 与相应的生产能耗 Y 的几组对照数据:

(1) 请画出散点图.

(2) 用最小二乘法求出 Y 关于 X 的线性回归方程 $\hat{y}=a+bx$. 并在散点图画出回归直线.

X	3	4	5	6
Y	2.5	3	4	4.5

12. 某数学老师身高 176 cm,他爷爷、父亲和儿子身高分别是 173 cm,170 cm,182 cm.因儿子的身高与父亲的身高有关,该老师用线性回归分析方法预测他孙子的身高为________ cm.

*13. 请推导线性回归一般公式,并证明回归方程的平移性质:

(1) 设散点图上有一系列点(x_i,y_i),$i=1,2,\cdots,n$,且 x_i 互异.其回归方程为 $\hat{y}=a+bx$. 请将 a,b 用 x_i,y_i 表示出来.

(2) 若散点均向右平移 c 单位长度为 (x_i+c, y_i), $i=1,2,\cdots,n$, 得到的回归方程为 $\hat{y}=a'+b'x$. 证明该回归直线即为原回归直线向右平移 c 单位长度所得到的直线 $\hat{y}=a+b(x-c)$. 即证明 $a'=a-bc$, $b'=b$.

14. 将以下二次曲线方程化成标准形

$$6x_1^2-4x_1x_2+3x_2^2-4x_1-8x_2+3=0.$$

15. 判断以下二次方程表示的是怎样的曲线

$$15x_1^2+6x_1x_2+11x_2^2-2x_1-4x_2+9=0.$$

部分习题参考答案

习 题 零

1. 因为 $A_{11}=a_{22}, A_{12}=-a_{21}, A_{21}=-a_{12}, A_{22}=a_{11}$，所以 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{22}=a_{11}A_{11}+a_{12}A_{12}=a_{21}A_{21}+a_{22}A_{22}=a_{11}A_{11}+a_{21}A_{21}=a_{12}A_{12}+a_{22}A_{22}$.

2. 证明性质 0.3′：不妨令第一行的元素可分成两个元素之和，

$$\begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$=(a_{11}+b_{11})\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}-(a_{12}+b_{12})\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}+(a_{13}+b_{13})\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$=a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}-a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}+a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}+b_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}-b_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

$$+b_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}+\begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

证明性质 0.4′：$\begin{vmatrix} ka_{11} & \cdots & ka_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}=\sum_{i=1}^{n}ka_{1i}A_{1i}=k\sum_{i=1}^{n}a_{1i}A_{1i}=k\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$.

若有一行为 0，记行列式为 D，将全是 0 的那行每个元素都乘 k，则 $D=kD$，所以 $D=0$.

证明性质 0.5′：假设令第 j 行 k 倍加到第 i 行，不妨设 $i<j$，

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+ka_{j1} & \cdots & a_{in}+ka_{jn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ ka_{j1} & \cdots & ka_{jn} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

3. 在 3 阶行列式中，若有两行元素相等，记该行列式为 D，交换元素相等的两行，得该行列式为 $-D$，则 $D=-D$，所以 $D=0$. 若两行元素成比例，则将倍数 k 提出行列式外，变为两行元素相等的情况.

4. (1) $a_{11}a_{22}a_{33}$.

(2) $\begin{vmatrix} 2 & -4 & -3 \\ 7 & 2 & 3 \\ 4 & -2 & 6 \end{vmatrix} = 2\begin{vmatrix} 2 & 3 \\ -2 & 6 \end{vmatrix} + 4\begin{vmatrix} 7 & 3 \\ 4 & 6 \end{vmatrix} - 3\begin{vmatrix} 7 & 2 \\ 4 & -2 \end{vmatrix} = 36+120+66=222.$

(3) $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix} = \begin{vmatrix} 3 & 4 \\ 4 & 5 \end{vmatrix} - 2\begin{vmatrix} 2 & 4 \\ 3 & 5 \end{vmatrix} + 3\begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} = -1+4-3=0.$

(4)
$$\begin{aligned} \begin{vmatrix} \lambda-2 & -3 & -2 \\ -1 & \lambda-8 & -2 \\ 2 & 14 & \lambda+3 \end{vmatrix} &= (\lambda-2)\begin{vmatrix} \lambda-8 & -2 \\ 14 & \lambda+3 \end{vmatrix} + 3\begin{vmatrix} -1 & -2 \\ 2 & \lambda+3 \end{vmatrix} - 2\begin{vmatrix} -1 & \lambda-8 \\ 2 & 14 \end{vmatrix} \\ &= (\lambda-2)(\lambda^2-5\lambda+4)+3(-\lambda+1)-2(2-2\lambda) \\ &= \lambda^3-7\lambda^2+15\lambda-9. \end{aligned}$$

5. (1) $\sum\limits_{i=2}^{n} i^2 = \sum\limits_{i=1}^{n} i^2 - 1 = \dfrac{n(n+1)(2n+1)}{6} - 1.$

$$\begin{aligned} \sum_{i=1}^{n}(i+1)(i+2) &= \sum_{i=1}^{n}(i^2+3i+2) = \sum_{i=1}^{n} i^2 + \sum_{i=1}^{n} 3i + \sum_{i=1}^{n} 2 \\ &= \frac{n(n+1)(2n+1)}{6} + \frac{3n(n+1)}{2} + 2n. \end{aligned}$$

(2) $\sum\limits_{i=1}^{n}(-1)^i = \begin{cases} -1, & n=2k+1, \\ 0, & n=2k, \end{cases} k \in \mathbb{Z}.$

$\sum\limits_{i=5}^{15}(-1)^i i = -10.$

6. $\sum\limits_{i=1}^{n} \dfrac{1}{i(i+1)} = \sum\limits_{i=1}^{n}\left(\dfrac{1}{i} - \dfrac{1}{i+1}\right) = 1 - \dfrac{1}{n+1} = \dfrac{n}{n+1}.$

7. $(a+b)^n = \sum\limits_{k=0}^{n} \mathrm{C}_n^k a^k b^{n-k}.$

8. (1) $\begin{cases} x=\dfrac{1}{4}, \\ y=-\dfrac{3}{4}, \\ z=\dfrac{3}{2}; \end{cases}$ (2) $\begin{cases} x=1, \\ y=1, \\ z=1. \end{cases}$

9. 若二元方程组有唯一解$\Leftrightarrow x=\dfrac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, y=\dfrac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \Leftrightarrow \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0.$

10. 首先,显然 $0,1 \in \mathbb{Q}(\mathrm{i})$.

其次,$\forall a+b\mathrm{i}, c+d\mathrm{i} \in \mathbb{Q}(\mathrm{i})$,

$$(a+b\mathrm{i}) \pm (c+d\mathrm{i}) = (a\pm c)+(b\pm d)\mathrm{i} \in \mathbb{Q}(\mathrm{i}),$$

$$(a+b\mathrm{i})(c+d\mathrm{i}) = (ac-bd)+(ad+bc)\mathrm{i} \in \mathbb{Q}(\mathrm{i}),$$

当 $c+d\mathrm{i} \neq 0$ 时,

$$\frac{a+b\mathrm{i}}{c+d\mathrm{i}} = \frac{1}{c^2+d^2}[(ac+bd)+(bc-ad)\mathrm{i}] \in Q(\mathrm{i}),$$

因此,$\mathbb{Q}(\mathrm{i}) = \{a+b\mathrm{i} \mid a,b \in \mathbb{Q}\}$ 是数域.

习 题 一

1. (1) $\boldsymbol{a} \perp \boldsymbol{b}$; (2) $\boldsymbol{a},\boldsymbol{b}$ 同向; (3) $\boldsymbol{a},\boldsymbol{b}$ 反向且 $|\boldsymbol{a}| \geqslant |\boldsymbol{b}|$; (4) $\boldsymbol{a},\boldsymbol{b}$ 同向.

2. (3)(6)成立,(1)(2)(4)(5)不成立.

3. $\dfrac{1}{2}\sqrt{\begin{vmatrix} y_2-y_1 & z_2-z_1 \\ y_3-y_1 & z_3-z_1 \end{vmatrix}^2 + \begin{vmatrix} z_2-z_1 & x_2-x_1 \\ z_3-z_1 & x_3-x_1 \end{vmatrix}^2 + \begin{vmatrix} x_2-x_1 & y_2-y_1 \\ x_3-x_1 & y_3-y_1 \end{vmatrix}^2}$.

4. $\dfrac{1}{6}\left|\begin{vmatrix} x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \\ x_4-x_1 & y_4-y_1 & z_4-z_1 \end{vmatrix}\right|$.

5. (1) $\boldsymbol{a} \cdot \boldsymbol{b} = 6, \langle \boldsymbol{a}, \boldsymbol{b} \rangle = \arccos \dfrac{1}{5}$;

(2) $(\boldsymbol{a})_{\boldsymbol{b}} = 6$, $\boldsymbol{b}$ 的方向余弦:$\dfrac{2}{3}, -\dfrac{1}{3}, \dfrac{2}{3}$.

6. $\boldsymbol{b}\times\boldsymbol{c} = (-\boldsymbol{a}-\boldsymbol{c})\times\boldsymbol{c} = -\boldsymbol{a}\times\boldsymbol{c}-\boldsymbol{c}\times\boldsymbol{c} = \boldsymbol{c}\times\boldsymbol{a}$,同理 $\boldsymbol{a}\times\boldsymbol{b} = \boldsymbol{b}\times\boldsymbol{c}$.

7. 由条件,

$$(\boldsymbol{a}\times\boldsymbol{b}) \cdot \boldsymbol{c} + (\boldsymbol{b}\times\boldsymbol{c}) \cdot \boldsymbol{c} + (\boldsymbol{c}\times\boldsymbol{a}) \cdot \boldsymbol{c} = 0,$$

因为$(\boldsymbol{b}\times\boldsymbol{c})\cdot\boldsymbol{c}=(\boldsymbol{c}\times\boldsymbol{a})\cdot\boldsymbol{c}=0$,

所以$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=0$,

从而 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 共面.

8. $x+y-3z-4=0$.

9. $2x+3y+z-6=0$.

10. $9y-z-2=0$.

11. $2x-4y-z+5=0$.

12. (1) 参数式:$\begin{cases}x=2+3t,\\y=-1-t,\\z=3+4t.\end{cases}$ 对称式:$\frac{x-2}{3}=\frac{y+1}{-1}=\frac{z-3}{4}$.

(2) 参数式:$\begin{cases}x=-t+2,\\y=3t+3,\\z=4t-5.\end{cases}$ 对称式:$\frac{x-2}{-1}=\frac{y-3}{3}=\frac{z+5}{4}$.

(3) 参数式:$\begin{cases}x=-4+5t,\\y=-1+4t,\\z=3-5t.\end{cases}$ 对称式:$\frac{x+4}{5}=\frac{y+1}{4}=\frac{z-3}{-5}$.

(4) 参数式:$\begin{cases}x=-3t,\\y=1+t,\\z=2+2t.\end{cases}$ 对称式:$\frac{x}{-3}=\frac{y-1}{1}=\frac{z-2}{2}$.

13. $d=\frac{6\sqrt{561}}{17}$,垂线方程$\frac{x-1}{14}=\frac{y-2}{-2}=\frac{z+5}{19}$.

14. (1) 由$\frac{x-5}{2}=\frac{y+3}{-2}=\frac{z-1}{3}$,令 $x=2t+5,y=-2t-3,z=3t+1$,代入 $x+2y-5z-11=0$,得 $2t+5+2(-2t-3)-5(3t+1)-11=0$,解得 $t=-1$,所以交点为$(3,-1,-2)$,直线与平面为相交关系.

(2) 相交,交点$\left(-\frac{7}{3},\frac{37}{3},-8\right)$.

15. l_1 与 l_2 交于$(1,-2,5)$,所以 l_1 与 l_2 共面.它们所在平面方程为 $2x-16y-13z+31=0$.

16. $60°$.

17. $0°$.

18. $\arccos\frac{\sqrt{2}}{3}$.

习　题　二

1. (1) $x_1=2,x_2=-1,x_3=1$;

（2）$x_1=1,x_2=-2,x_3=3$；

（3）$x_1=2,x_2=-1,x_3=1,x_4=-3$；

（4）$x_1=5,x_2=-2,x_3=1$；

（5）$x_1=-8,x_2=3,x_3=6,x_4=0$；

（6）无解.

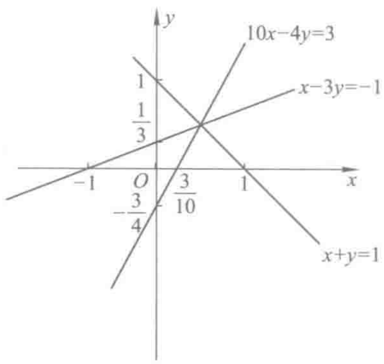
2. $a=-1$ 时，方程组有解.

一般解为 $\begin{cases} x_1=\dfrac{1}{7}(1-18x_3), \\ x_2=\dfrac{1}{7}(2-x_3), \end{cases}$ 其中 x_3 为自由未知量.

3. $a=-\dfrac{2}{3}$时，无解；$a\neq-\dfrac{2}{3}$时，有唯一解.

4.（1）有唯一解 $\begin{cases} x=\dfrac{1}{2}, \\ y=\dfrac{1}{2}; \end{cases}$

（2）



5. $a=-2$ 时，方程组有解.

一般解为 $\begin{cases} x_1=-3x_3-2, \\ x_2=2x_3+5, \\ x_4=-10, \end{cases}$ 其中 x_3 为自由未知量.

6.（1）$\begin{cases} x_1=x_3, \\ x_2=2x_3, \\ x_4=-3x_3, \end{cases}$ x_3 为自由未知量；

（2）$\begin{cases} x_1=\dfrac{55}{41}x_4, \\ x_2=\dfrac{10}{41}x_4，x_4 \text{ 为自由未知量}. \\ x_3=-\dfrac{33}{41}x_4, \end{cases}$

习 题 三

1. （1）$\dfrac{(n-1)(n-2)}{2}$；（2）$n-1$.

2. $\dfrac{n(n-1)}{2}-r$.

3. （1）−49；（2）103；（3）155；（4）160.

4. （1）$(a-1)^n+n(a-1)^{n-1}$.

（2）$n=1$ 时，$|a_1+b_1|=a_1+b_1$；

$n=2$ 时，$\begin{vmatrix} a_1+b_1 & a_1+b_2 \\ a_2+b_1 & a_2+b_2 \end{vmatrix}=(a_2-a_1)(b_1-b_2)$；

$n\geqslant 3$ 时，$\begin{vmatrix} a_1+b_1 & a_1+b_2 & \cdots & a_1+b_n \\ a_2+b_1 & a_2+b_2 & \cdots & a_2+b_n \\ \vdots & \vdots & & \vdots \\ a_n+b_1 & a_n+b_2 & \cdots & a_n+b_n \end{vmatrix}=0.$

5. 记 $A(x,y,z)=\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$,

$$\begin{aligned} \text{左}&=A(x,y,z)+A(x,y,x)+A(x,z,z)+A(x,z,x) \\ &\quad +A(y,y,z)+A(y,y,x)+A(y,z,z)+A(y,z,x) \\ &=A(x,y,z)+A(y,z,x) \\ &=2A(x,y,z)=\text{右}. \end{aligned}$$

6. $\begin{vmatrix} x & x^2 & yz \\ y & y^2 & xz \\ z & z^2 & xy \end{vmatrix}=\begin{vmatrix} x & x^2 & yz \\ y-x & y^2-x^2 & z(x-y) \\ z-x & z^2-x^2 & y(x-z) \end{vmatrix}=(x-y)(z-x)\begin{vmatrix} x & x^2 & yz \\ -1 & -(x+y) & z \\ 1 & z+x & -y \end{vmatrix}$

$$=(x-y)(z-x)\begin{vmatrix} 0 & -xz & y(z+x) \\ 0 & z-y & z-y \\ 1 & z+x & -y \end{vmatrix}$$

$$=(x-y)(y-z)(z-x)\begin{vmatrix} -xz & y(z+x) \\ -1 & -1 \end{vmatrix}$$

$$=(xy+yz+zx)(x-y)(y-z)(z-x).$$

7. (1) 27; (2) 0; (3) −45.

8. 必要性.

因为方程组$\begin{cases} ax+2by+3c=0, \\ bx+2cy+3a=0, \\ cx+2ay+3b=0 \end{cases}$有唯一解$(x_0,y_0)$,

所以关于 x,y,z 的齐次方程组$\begin{cases} ax+2by+3cz=0, \\ bx+2cy+3az=0, \\ cx+2ay+3bz=0 \end{cases}$有非零解:$(x_0,y_0,1)$.

根据定理 3.5 知 $\begin{vmatrix} a & 2b & 3c \\ b & 2c & 3a \\ c & 2a & 3b \end{vmatrix}=0 \Rightarrow \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}=0$

$$\Rightarrow (a+b+c)(a^2+b^2+c^2-ab-bc-ca)=0.$$

因为 a,b,c 不全相等(否则三条直线重合),

所以 $a^2+b^2+c^2-ab-bc-ca=\frac{1}{2}[(a-b)^2+(b-c)^2+(c-a)^2]\neq 0$,

从而 $a+b+c=0$.

充分性.

若 $a+b+c=0$,则$\begin{cases} ax+2by+3c=0, \\ bx+2cy+3a=0, \\ cx+2ay+3b=0 \end{cases} \Leftrightarrow \begin{cases} ax+2by+3c=0, \\ bx+2cy+3a=0. \end{cases}$

而 $b^2=(a+c)^2\neq ac$(否则 $a^2+ac+c^2=0 \Rightarrow a=c=0 \Rightarrow a=b=c=0$ 与已知矛盾!),

所以系数矩阵的行列式$\begin{vmatrix} a & 2b \\ b & 2c \end{vmatrix}=2(ac-b^2)\neq 0$.

故原方程组有唯一解⇔三条直线交于一点.

9. $a\neq 1$ 且 $b\neq 0$.

10. $a\neq 1$ 且 $b\neq 0$ 时,有唯一解;
 $b=0$ 时,有无穷多解;
 $a=1$ 且 $b\neq 0$ 时,无解.

习 题 四

1. (1)(2)不成立,(3)成立.

2. $\begin{pmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \end{pmatrix}$.

3. $x=2, y=1$.

4. (1) $\begin{pmatrix} 3 & -3 \\ -5 & -7 \\ 9 & 15 \end{pmatrix}$； (2) $\begin{pmatrix} 1 & 0 & 5 & 3 \\ 0 & -1 & 0 & 0 \end{pmatrix}$；

(3) $a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2+2a_{12}x_1x_2+2a_{13}x_1x_3+2a_{23}x_2x_3$；

(4) $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$.

5. 略.

6. $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$, a, b 为自由变量.

7. (1) $\begin{pmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{7}{2} & \frac{1}{2} \end{pmatrix}$；

(2) $\begin{pmatrix} -3 & -2 & -\frac{5}{2} \\ -16 & -11 & -13 \\ -12 & -7 & -\frac{19}{2} \end{pmatrix}$.

8. $\boldsymbol{A}^2-2\boldsymbol{A}+5\boldsymbol{I}=\boldsymbol{O}$

$\Rightarrow \boldsymbol{A}(\boldsymbol{A}-2\boldsymbol{I})=-5\boldsymbol{I}$

$\Rightarrow \boldsymbol{A}\left(\frac{2}{5}\boldsymbol{I}-\frac{1}{5}\boldsymbol{A}\right)=\boldsymbol{I}$

$\Rightarrow \boldsymbol{A}$ 可逆，且 $\boldsymbol{A}^{-1}=\frac{2}{5}\boldsymbol{I}-\frac{1}{5}\boldsymbol{A}$.

9. (1) $\boldsymbol{A}^{-1}=\begin{pmatrix} 2 & 1 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 2 & 3 \end{pmatrix}$； (2) $\boldsymbol{A}^{-1}=\begin{pmatrix} 0 & 0 & \frac{3}{5} & -\frac{1}{5} \\ 0 & 0 & -\frac{1}{5} & \frac{2}{5} \\ -\frac{1}{3} & \frac{2}{3} & 0 & 0 \\ \frac{2}{3} & -\frac{1}{3} & 0 & 0 \end{pmatrix}$.

10. (1) 在 $2\boldsymbol{A}^{-1}\boldsymbol{B}=\boldsymbol{B}-4\boldsymbol{I}$ 两边同时左乘 $\boldsymbol{A}$，得

$2\boldsymbol{B}=\boldsymbol{AB}-4\boldsymbol{A}$,

所以 $\boldsymbol{AB}-2\boldsymbol{B}=4\boldsymbol{A}$，即 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{B}=4\boldsymbol{A}$. 两边同时右乘 $\boldsymbol{A}^{-1}$，得 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{BA}^{-1}=4\boldsymbol{I}$，所以 $(\boldsymbol{A}-2\boldsymbol{I})\frac{\boldsymbol{BA}^{-1}}{4}=\boldsymbol{I}$.

从而 $\boldsymbol{A}-2\boldsymbol{I}$ 可逆.

(2) 因为 $\boldsymbol{B}=\begin{pmatrix}1 & -2 & 0\\ 1 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}$可逆,且 $\boldsymbol{B}^{-1}=\begin{pmatrix}\frac{1}{2} & \frac{1}{2} & 0\\ -\frac{1}{4} & \frac{1}{4} & 0\\ 0 & 0 & \frac{1}{2}\end{pmatrix}$.

所以,由 $2\boldsymbol{A}^{-1}\boldsymbol{B}=\boldsymbol{B}-4\boldsymbol{I}$ 可得$(\boldsymbol{I}-2\boldsymbol{A}^{-1})\boldsymbol{B}=4\boldsymbol{I}$.

从而$\dfrac{\boldsymbol{I}-2\boldsymbol{A}^{-1}}{4}=\boldsymbol{B}^{-1}$.所以,

$$\boldsymbol{I}-2\boldsymbol{A}^{-1}=4\boldsymbol{B}^{-1}=\begin{pmatrix}2 & 2 & 0\\ -1 & 1 & 0\\ 0 & 0 & 2\end{pmatrix}$$

所以,$\boldsymbol{A}^{-1}=\dfrac{\boldsymbol{I}-4\boldsymbol{B}^{-1}}{2}=\dfrac{1}{2}\begin{pmatrix}-1 & -2 & 0\\ 1 & 0 & 0\\ 0 & 0 & -1\end{pmatrix}=\begin{pmatrix}-\frac{1}{2} & -1 & 0\\ \frac{1}{2} & 0 & 0\\ 0 & 0 & -\frac{1}{2}\end{pmatrix}$.

从而 $\boldsymbol{A}=\begin{pmatrix}0 & 2 & 0\\ -1 & -1 & 0\\ 0 & 0 & -2\end{pmatrix}$.

11. $\boldsymbol{A}^2=\begin{pmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{pmatrix}$,$\boldsymbol{A}^4=\boldsymbol{I}$.

所以 $\boldsymbol{B}^{2\,016}=\boldsymbol{P}^{-1}\boldsymbol{A}^{2\,016}\boldsymbol{P}=\boldsymbol{P}^{-1}\boldsymbol{I}\boldsymbol{P}=\boldsymbol{I}$,

从而 $\boldsymbol{B}^{2\,016}-2\boldsymbol{A}^2=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}-\begin{pmatrix}-2 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & 2\end{pmatrix}=\begin{pmatrix}3 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & -1\end{pmatrix}$.

习 题 五

1. (1) 线性无关; (2) 线性相关.
2. $k\neq 0$ 且 $k\neq 1$.
3. 因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性相关,

   所以 $\exists k_1,k_2,k_3,k_4$,使得 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\alpha}_3+k_4\boldsymbol{\alpha}_4=\boldsymbol{0}$.

下证 k_1,k_2,k_3,k_4 全不为 0. 假设 k_1,k_2,k_3,k_4 中有一个为 0.

不妨设为 k_1，则 $k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\alpha}_3+k_4\boldsymbol{\alpha}_4=\boldsymbol{0}$，且 k_2,k_3,k_4 不全为 0. 与 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关矛盾.

因此，结论成立.

4. $\lambda=1$.

5. (1) 秩 $=3$，极大线性无关组：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4$（答案不唯一）.

   (2) 秩 $=3$.极大线性无关组：$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$.

6. 因为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 线性表出，

   由定理 5.3，$r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)\leqslant r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)\leqslant 2$，从而 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性相关.

7. 1°，2°，由矩阵的秩与 k 阶子式的定义易证；

   3°记 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$，$\boldsymbol{B}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$. 不妨设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 为 $\boldsymbol{A}$ 的列向量组的极大线性无关组，而 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 为 $\boldsymbol{B}$ 的列向量组的极大线性无关组，所以 $\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2,\cdots,\boldsymbol{\alpha}_n+\boldsymbol{\beta}_n$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性表示，则由定理 5.3 知 $r(\boldsymbol{A}+\boldsymbol{B})\leqslant r+s=r(\boldsymbol{A})+r(\boldsymbol{B})$.

   4°考虑矩阵 $\boldsymbol{A},\boldsymbol{B}$ 与 $(\boldsymbol{A},\boldsymbol{B})$ 的列秩，即得结论.

   5°记 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$，$\boldsymbol{AB}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s)$，$\boldsymbol{B}=(b_{ij})_{n\times s}$. 则

$$\begin{aligned}\boldsymbol{\beta}_1&=b_{11}\boldsymbol{\alpha}_1+b_{21}\boldsymbol{\alpha}_2+\cdots+b_{n1}\boldsymbol{\alpha}_n,\\ \boldsymbol{\beta}_2&=b_{12}\boldsymbol{\alpha}_1+b_{22}\boldsymbol{\alpha}_2+\cdots+b_{n2}\boldsymbol{\alpha}_n,\\ &\cdots\cdots\cdots\cdots\\ \boldsymbol{\beta}_s&=b_{1s}\boldsymbol{\alpha}_1+b_{2s}\boldsymbol{\alpha}_2+\cdots+b_{ns}\boldsymbol{\alpha}_n,\end{aligned}$$

   从而，$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s\}$ 可由 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 线性表出，由定理 5.3 有

$$r(\boldsymbol{A})=r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\geqslant r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s)=r(\boldsymbol{AB}).$$

   同理可得 $r(\boldsymbol{B})\geqslant r(\boldsymbol{AB})$.

   6°由高斯消元法 $\boldsymbol{A}$ 通过初等行变换可以化为阶梯形矩阵，其非零行的行数为 $r(\boldsymbol{A})$，$\boldsymbol{B}$ 通过初等行变换化为阶梯形矩阵，其非零行的行数为 $r(\boldsymbol{B})$，则 $\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{B}\end{pmatrix}$ 通过初等行变换可以化为阶梯形矩阵，其非零行的行数为 $r(\boldsymbol{A})+r(\boldsymbol{B})$，故

$$r\left(\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{B}\end{pmatrix}\right)=r(\boldsymbol{A})+r(\boldsymbol{B}).$$

   7° $\boldsymbol{A}$ 有一个 $r(\boldsymbol{A})$ 阶子式不为零，$\boldsymbol{B}$ 有一个 $r(\boldsymbol{B})$ 阶子式不为零，故 $\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\ \boldsymbol{C}&\boldsymbol{B}\end{pmatrix}$ 必有一个 $r(\boldsymbol{A})+r(\boldsymbol{B})$ 阶子式不为零.

   8°设 $r(\boldsymbol{A})=r$，则有可逆矩阵 $\boldsymbol{P}$，$\boldsymbol{Q}$ 使 $\boldsymbol{PAQ}=\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{O}\end{pmatrix}$，于是

$$\boldsymbol{PAB}=\boldsymbol{PAQQ}^{-1}\boldsymbol{B}=\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{O}\end{pmatrix}\boldsymbol{Q}^{-1}\boldsymbol{B},$$

对 $\boldsymbol{Q}^{-1}\boldsymbol{B}$ 作相应的分块，前 r 行分为一块，后 $n-r$ 行分为一块，记为 $\boldsymbol{Q}^{-1}\boldsymbol{B}=\begin{pmatrix}\boldsymbol{C}_1\\ \boldsymbol{C}_2\end{pmatrix}$，

则

$$\boldsymbol{PAB}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\begin{pmatrix}\boldsymbol{C}_1\\ \boldsymbol{C}_2\end{pmatrix}=\begin{bmatrix}\boldsymbol{C}_1\\ \boldsymbol{O}\end{bmatrix}.$$

由 $\boldsymbol{P}$ 可逆，得

$$r(\boldsymbol{AB})=r((\boldsymbol{AB})^{\mathrm{T}})=r(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}})\leqslant r(\boldsymbol{B}^{\mathrm{T}})=r(\boldsymbol{B}).$$

综上即得结论.

9°因为$\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{I} & \boldsymbol{B}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{O} & -\boldsymbol{AB}\\ \boldsymbol{I} & \boldsymbol{B}\end{pmatrix}=\begin{pmatrix}\boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{I} & \boldsymbol{B}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{I} & \boldsymbol{O}\end{pmatrix}$，

由于初等变换不改变矩阵行秩和列秩，所以 $r\left(\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{I} & \boldsymbol{B}\end{pmatrix}\right)=r\left(\begin{pmatrix}\boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{I} & \boldsymbol{O}\end{pmatrix}\right)=n$，

又由性质 7°，有 $r(\boldsymbol{A})+r(\boldsymbol{B})\leqslant r\left(\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{I} & \boldsymbol{B}\end{pmatrix}\right)=n.$

8. $\boldsymbol{\eta}_1=\begin{pmatrix}4\\ -9\\ 4\\ 3\end{pmatrix}$.

9. 设 $\boldsymbol{B}$ 的列向量为 $\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_n$.
   由 $\boldsymbol{AB}=\boldsymbol{O}$ 得 $\boldsymbol{AB}_i=\boldsymbol{O}$.
   $\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_n$ 是齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解.
   由定理 5.8，$r(\boldsymbol{B})=r(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)\leqslant n-r(\boldsymbol{A})$，
   即 $r(\boldsymbol{A})+r(\boldsymbol{B})\leqslant n$.

10. (1) 4； (2) 3.

11. 由 $\boldsymbol{AB}=\boldsymbol{O}$，设 $\boldsymbol{B}$ 的列向量为 $\boldsymbol{B}_1,\boldsymbol{B}_2,\boldsymbol{B}_3$，因为 $\boldsymbol{B}$ 是 3 阶非零矩阵，所以 $r(\boldsymbol{B})\geqslant 1$，$\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解，所以 $|\boldsymbol{A}|=0$，由 $|\boldsymbol{A}|=\begin{vmatrix}1 & 2 & -1\\ 3 & -1 & 0\\ 2 & x & 1\end{vmatrix}=0$，得 $x=-3$，所以 $r(\boldsymbol{A})=2$，又 $r(\boldsymbol{A})+r(\boldsymbol{B})\leqslant 3$，所以 $r(\boldsymbol{B})\leqslant 1$，$r(\boldsymbol{B})=1$.

12. (1) 相应的齐次线性方程组有 2 个线性无关的解，
    所以 $2=4-r(\boldsymbol{A})$，$r(\boldsymbol{A})=2$.

    (2) $a=2$，$b=-3$.

    通解：$\begin{pmatrix}x_1\\ x_2\\ x_3\\ x_4\end{pmatrix}=\begin{pmatrix}2\\ -3\\ 0\\ 0\end{pmatrix}+k_1\begin{pmatrix}-2\\ 1\\ 1\\ 0\end{pmatrix}+k_2\begin{pmatrix}4\\ -5\\ 0\\ 1\end{pmatrix}$.

13. 对 $\forall \lambda_1,\cdots,\lambda_n \in \mathbb{R}^n$,

$$(\boldsymbol{\beta},\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_n\boldsymbol{\alpha}_n)=\sum_{i=1}^{n}\lambda_i(\boldsymbol{\beta},\boldsymbol{\alpha}_i)=0.$$

习 题 六

1. (1) $\lambda_1=1$(两重), $\lambda_2=10$.

属于 λ_1 的特征向量:$k_1\begin{pmatrix}-2\\1\\0\end{pmatrix}+k_2\begin{pmatrix}2\\0\\1\end{pmatrix}$, k_1,k_2 不全为 0.

属于 λ_2 的特征向量:$k_3\begin{pmatrix}1\\2\\-2\end{pmatrix}$, $k_3\neq0$.

(2) $\lambda_1=1$, $\lambda_2=3$(两重).

属于 λ_1 的特征向量:$k_1\begin{pmatrix}-2\\0\\1\end{pmatrix}$, $k_1\neq0$.

属于 λ_2 的特征向量:$k_2\begin{pmatrix}-1\\1\\-2\end{pmatrix}$, $k_2\neq0$.

(3) $\lambda=-1$(三重).

特征向量:$k\begin{pmatrix}-1\\-1\\1\end{pmatrix}$, $k\neq0$.

(4) $\lambda_1=0$, $\lambda_2=1$, $\lambda_3=-1$.

属于 $\lambda_1,\lambda_2,\lambda_3$ 的特征向量分别是 $k_1\begin{pmatrix}1\\1\\-1\end{pmatrix}$, $k_2\begin{pmatrix}1\\1\\1\end{pmatrix}$, $k_3\begin{pmatrix}1\\-1\\-1\end{pmatrix}$, k_1,k_2,k_3 均不为 0.

2. 提示:$\boldsymbol{A}(2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)=2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$.

3. $\begin{pmatrix}2&1\\1&1\end{pmatrix}^{-1}\boldsymbol{A}\begin{pmatrix}2&1\\1&1\end{pmatrix}=\begin{pmatrix}2&0\\0&3\end{pmatrix}$,

$$\boldsymbol{A}^m=\begin{pmatrix}2&1\\1&1\end{pmatrix}\begin{pmatrix}2&0\\0&3\end{pmatrix}^m\begin{pmatrix}2&1\\1&1\end{pmatrix}^{-1}=\begin{pmatrix}2^{m+1}-3^m & -2^{m+1}+2\cdot3^m\\2^m-3^m & -2^m+2\cdot3^m\end{pmatrix}.$$

4. (1) $r(\boldsymbol{A})=2<3$,所以 $|\boldsymbol{A}|=0$,

从而 0 是 $\boldsymbol{A}$ 的特征值,对应的特征向量与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 正交,求得为$(1,-1,-1)^{\mathrm{T}}$.

(2) 设 $\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta})$,其中 $\boldsymbol{\beta}=(1,-1,-1)^{\mathrm{T}}$.

则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(6,6,0)$,$\boldsymbol{A}=\boldsymbol{P}\mathrm{diag}(6,6,0)\boldsymbol{P}^{-1}=\begin{pmatrix}4&2&2\\2&4&-2\\2&-2&4\end{pmatrix}$.

5. (1) 因为 $\boldsymbol{A}\sim\boldsymbol{B}$,所以 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=|\lambda\boldsymbol{I}-\boldsymbol{B}|$,解得 $x=0,y=-2$;

(2) 仿照例 6.3,$\boldsymbol{P}=\begin{pmatrix}0&0&1\\-2&1&0\\1&1&-1\end{pmatrix}$(答案不唯一).

6. (1) $\boldsymbol{B}=\begin{pmatrix}1&0&0\\1&2&2\\1&1&3\end{pmatrix}$;

(2) $\boldsymbol{A}\sim\boldsymbol{B}\Rightarrow\boldsymbol{A}$ 与 $\boldsymbol{B}$ 特征值相同,为 $\lambda_1=1$(两重),$\lambda_2=4$.

7. (1) $\boldsymbol{AP}=(\boldsymbol{AX},\boldsymbol{A}^2\boldsymbol{X},\boldsymbol{A}^3\boldsymbol{X})=(\boldsymbol{X},\boldsymbol{AX},\boldsymbol{A}^2\boldsymbol{X})\begin{pmatrix}0&0&0\\1&0&3\\0&1&-2\end{pmatrix}$,

所以 $\boldsymbol{B}=\begin{pmatrix}0&0&0\\1&0&3\\0&1&-2\end{pmatrix}$;

(2) $|\boldsymbol{A}+\boldsymbol{I}|=|\boldsymbol{P}(\boldsymbol{B}+\boldsymbol{I})\boldsymbol{P}^{-1}|=|\boldsymbol{P}|\,|\boldsymbol{B}+\boldsymbol{I}|\,|\boldsymbol{P}^{-1}|=|\boldsymbol{B}+\boldsymbol{I}|=-4$.

8. (1) $\lambda=0,a=-2,b=-1$;

(2) $\boldsymbol{A}$ 有 3 个不同的特征值:$0,-1\pm2\mathrm{i}$,

所以 $\boldsymbol{A}$ 可对角化.

9. (1) 因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是 $\boldsymbol{AX}=\boldsymbol{0}$ 的解,

所以 $\boldsymbol{A\alpha}_1=\boldsymbol{0},\boldsymbol{A\alpha}_2=\boldsymbol{0}$.

从而 0 是 $\boldsymbol{A}$ 的特征值,相应特征向量是 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2$($k_1,k_2$ 不同时为 0).

又因为 $\boldsymbol{A}$ 各行元素之和为 3,

所以 $\boldsymbol{A}(1,1,1)^{\mathrm{T}}=3(1,1,1)^{\mathrm{T}}$,

从而 3 是 $\boldsymbol{A}$ 的特征值,相应特征向量是 $k_3(1,1,1)^{\mathrm{T}}(k_3\neq0)$.

(2) $\boldsymbol{\Lambda}=\mathrm{diag}(0,0,3)$.

$\boldsymbol{Q}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,(1,1,1)^{\mathrm{T}}$ 正交化得到

$$\boldsymbol{Q}=\begin{pmatrix}\frac{\sqrt{6}}{3}&0&\frac{\sqrt{3}}{3}\\-\frac{\sqrt{6}}{6}&-\frac{\sqrt{2}}{2}&\frac{\sqrt{3}}{3}\\-\frac{\sqrt{6}}{6}&\frac{\sqrt{2}}{2}&\frac{\sqrt{3}}{3}\end{pmatrix}\text{(答案不唯一)}.$$

(3) $\boldsymbol{A}=\boldsymbol{P\Lambda P}^{-1}$,其中 $\boldsymbol{P}=\begin{pmatrix}-1&0&1\\2&-1&1\\-1&1&1\end{pmatrix}$.$\boldsymbol{A}=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}$.

$\left(\boldsymbol{A}-\frac{3}{2}\boldsymbol{I}\right)^6=\boldsymbol{P}\left(\boldsymbol{\Lambda}-\frac{3}{2}\boldsymbol{I}\right)^6\boldsymbol{P}^{-1}=\left(\frac{3}{2}\right)^6\boldsymbol{I}$.

习 题 七

1. (1) 二次型 $f(x_1,x_2,x_3)=x_1^2+2x_2^2+2x_1x_2-2x_1x_3$ 的矩阵为 $\boldsymbol{A}=\begin{pmatrix}1&1&-1\\1&2&0\\-1&0&0\end{pmatrix}$.下面求一矩阵 $\boldsymbol{C}$,使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ 为对角矩阵.

$$\begin{pmatrix}1&1&-1\\1&2&0\\-1&0&0\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0\\1&1&1\\-1&1&-1\\1&-1&1\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&1&1\\0&1&-1\\1&-1&1\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\\1&-1&2\\0&1&-1\\0&0&1\end{pmatrix},$$

所以 $\boldsymbol{C}=\begin{pmatrix}1&-1&2\\0&1&-1\\0&0&1\end{pmatrix}$,且 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix}$.

从而 $f(x_1,x_2,x_3)$ 的一个标准形为 $y_1^2+y_2^2-2y_3^2$.非退化线性替换为 $\boldsymbol{X}=\boldsymbol{CY}$.

(2) 二次型 $f(x_1,x_2,x_3)=x_1^2-x_3^2+2x_1x_2+2x_2x_3$ 的矩阵为 $\boldsymbol{A}=\begin{pmatrix}1&1&0\\1&0&1\\0&1&-1\end{pmatrix}$.

下面求一矩阵 $\boldsymbol{C}$,使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ 为对角矩阵.

$$\begin{pmatrix}1&1&0\\1&0&1\\0&1&-1\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0\\1&-1&1\\0&1&-1\\1&-1&0\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&-1&1\\0&1&-1\\1&-1&0\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\\1&-1&-1\\0&1&1\\0&0&1\end{pmatrix},$$

所以 $C=\begin{pmatrix}1&-1&-1\\0&1&1\\0&0&1\end{pmatrix}$,且 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix}$.

从而 $f(x_1,x_2,x_3)$ 的一个标准形为 $y_1^2-y_2^2$.非退化线性替换为 $\boldsymbol{X}=\boldsymbol{CY}$.

(3) 二次型 $f(x_1,x_2,x_3)=x_1x_2+x_1x_3+x_2x_3$ 的矩阵为 $\boldsymbol{A}=\begin{pmatrix}0&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&0&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&0\end{pmatrix}$.

下面求一矩阵 $\boldsymbol{C}$,使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ 为对角矩阵.

$$\begin{pmatrix}0&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&0&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&0\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\\1&0&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&0\\1&0&0\\0&1&0\\1&0&1\end{pmatrix}\to\begin{pmatrix}1&1&\frac{1}{2}\\1&0&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&0\\1&0&0\\0&1&0\\1&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0\\1&-1&0\\\frac{1}{2}&0&-\frac{1}{4}\\1&-1&-\frac{1}{2}\\0&1&0\\1&-1&\frac{1}{2}\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-\frac{1}{4}\\1&-1&-\frac{1}{2}\\0&1&0\\1&-1&\frac{1}{2}\end{pmatrix},$$

所以 $\boldsymbol{C}=\begin{pmatrix}1&-1&-\frac{1}{2}\\0&1&0\\1&-1&\frac{1}{2}\end{pmatrix}$,且 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-\frac{1}{4}\end{pmatrix}$.

从而 $f(x_1,x_2,x_3)$ 的一个标准形为 $y_1^2-y_2^2-\frac{1}{4}y_3^2$.非退化线性替换为 $\boldsymbol{X}=\boldsymbol{CY}$.

(4) $f(x_1,x_2,x_3,x_4)=2x_1x_2-2x_3x_4$ 的矩阵为 $\boldsymbol{A}=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&-1&0\end{pmatrix}$.

下面求一矩阵 $\boldsymbol{C}$,使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ 为一对角矩阵.

$$\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&-1&0\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\to\begin{pmatrix}2&1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&-1&0\\1&0&0&0\\1&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}$$

$$\to\begin{pmatrix}2&0&0&0\\0&-\frac{1}{2}&0&0\\0&0&0&-1\\0&0&-1&0\\1&-\frac{1}{2}&0&0\\1&\frac{1}{2}&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\to\begin{pmatrix}2&0&0&0\\0&-\frac{1}{2}&0&0\\0&0&-2&0\\0&0&0&\frac{1}{2}\\1&-\frac{1}{2}&0&0\\1&\frac{1}{2}&0&0\\0&0&1&-\frac{1}{2}\\0&0&1&\frac{1}{2}\end{pmatrix},$$

所以 $C=\begin{pmatrix}1&-\frac{1}{2}&0&0\\1&\frac{1}{2}&0&0\\0&0&1&-\frac{1}{2}\\0&0&1&\frac{1}{2}\end{pmatrix}$，且 $C^{\mathrm{T}}AC=\begin{pmatrix}2&&&\boldsymbol{O}\\&-\frac{1}{2}&&\\&&-2&\\\boldsymbol{O}&&&\frac{1}{2}\end{pmatrix}$

从而 $f(x_1,x_2,x_3,x_4)$ 的一个标准形为 $2y_1^2-\frac{1}{2}y_2^2-2y_3^2+\frac{1}{2}y_4^2$.非退化线性替换为 $X=CY$.

2. (1) $A=\begin{pmatrix}2&2&-2\\2&5&-4\\-2&-4&5\end{pmatrix}$，特征多项式为 $|\lambda I-A|=\begin{vmatrix}\lambda-2&-2&2\\-2&\lambda-5&4\\2&4&\lambda-5\end{vmatrix}=$ $(\lambda-1)^2(\lambda-10)$.

所以 A 的特征值为 $\lambda_1=1$（二重），$\lambda_2=10$.

对于特征值 1，求解齐次线性方程组 $(I-A)X=\mathbf{0}$，

得两个线性无关的单位正交特征向量 $\boldsymbol{\mu}_1=\left(\frac{2\sqrt{5}}{5},0,\frac{\sqrt{5}}{5}\right)^{\mathrm{T}}$，$\boldsymbol{\mu}_2=\left(-\frac{2\sqrt{5}}{15},\frac{\sqrt{5}}{3},\frac{4\sqrt{5}}{15}\right)^{\mathrm{T}}$.

对于特征值 10，求解齐次线性方程组 $(10I-A)X=\mathbf{0}$，

得出一个单位特征向量 $\boldsymbol{\mu}_3=\left(\frac{1}{3},\frac{2}{3},-\frac{2}{3}\right)^{\mathrm{T}}$.

令 $\boldsymbol{Q}=\begin{pmatrix} \frac{2\sqrt{5}}{5} & -\frac{2\sqrt{5}}{15} & \frac{1}{3} \\ 0 & \frac{\sqrt{5}}{3} & \frac{2}{3} \\ \frac{\sqrt{5}}{5} & \frac{4\sqrt{5}}{15} & -\frac{2}{3} \end{pmatrix}$，则 $\boldsymbol{Q}$ 为正交矩阵，且 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix}$.

所以 $f(x_1,x_2,x_3)$ 的一个标准形为 $y_1^2+y_2^2+10y_3^2$.所作正交替换为 $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Y}$.

（2）$\boldsymbol{A}=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$，特征多项式为

$$f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix} \lambda & -1 & 0 & 0 \\ -1 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 1 & \lambda \end{vmatrix}=(\lambda+1)^2(\lambda-1)^2.$$

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=1$（二重），$\lambda_2=-1$（二重）.

对于特征值 $\lambda=1$，求解方程组 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得两个线性无关的单位特征向量

$\boldsymbol{\mu}_1=\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},0,0\right)^{\mathrm{T}}$，$\boldsymbol{\mu}_2=\left(0,0,\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)^{\mathrm{T}}$.

对于特征值 $\lambda=-1$，求解方程组 $(-\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得两个线性无关的单位特征向量

$\boldsymbol{\mu}_3=\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2},0,0\right)^{\mathrm{T}}$，$\boldsymbol{\mu}_4=\left(0,0,\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)^{\mathrm{T}}$.

令 $\boldsymbol{Q}=\begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{0} & 0 & -\frac{\sqrt{2}}{2} & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix}$，则 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

所以 $f(x_1,x_2,x_3,x_4)$ 的一个标准形为 $y_1^2+y_2^2-y_3^2-y_4^2$.所作正交替换为 $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{Y}$.

3. $\boldsymbol{A}=\begin{pmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{pmatrix}$，特征多项式为 $f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix} \lambda-2 & 0 & -4 \\ 0 & \lambda-6 & 0 \\ -4 & 0 & \lambda-2 \end{vmatrix}=(\lambda+2)(\lambda-6)^2$.

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=-2$，$\lambda_2=6$（二重）.

对于特征值 $\lambda=-2$，求解方程组$(-2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得到一个单位特征向量
$\boldsymbol{\mu}_1=\left(-\frac{\sqrt{2}}{2},0,\frac{\sqrt{2}}{2}\right)^{\mathrm{T}}$.

对于特征值 $\lambda=6$，求解方程组$(6\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得到两个线性无关的单位特征向量
$\boldsymbol{\mu}_2=(0,1,0)^{\mathrm{T}},\boldsymbol{\mu}_3=\left(\frac{\sqrt{2}}{2},0,\frac{\sqrt{2}}{2}\right)^{\mathrm{T}}$.

$$令\ \boldsymbol{Q}=\begin{pmatrix}-\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2}\\ 0 & 1 & 0\\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2}\end{pmatrix}，则\ \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}-2 & & \\ & 6 & \\ & & 6\end{pmatrix}.$$

所以 $2x^2+6y^2+2z^2+8xz-1=0$ 的标准方程为 $6x_1^2+6y_1^2-2z_1^2=1$，为单叶双曲面.

4. 必要性显然.下证充分性：

若 $\forall\,\boldsymbol{\alpha}\in F^n,\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=0$，则

① 取 $\boldsymbol{\alpha}_1=(0,\cdots,0,1,0,\cdots,0)^{\mathrm{T}}$，1 在第 i 个位置.

记 $\boldsymbol{A}=(a_{ij})_{n\times n}$，则 $\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}_i=a_{ii}=0,\ \forall\, i=1,2,\cdots,n$.

② 取 $\boldsymbol{\alpha}_{ij}=(0,\cdots,0,1,0,\cdots,0,1,0,\cdots,0)^{\mathrm{T}}$，1 处于第 i 和第 j 个位置$(i\neq j)$，则 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}=a_{ij}+a_{ji}=0$. 由于 $\boldsymbol{A}$ 对称，所以 $a_{ij}=a_{ji}=0$.

综上所述，$\boldsymbol{A}=\boldsymbol{O}$.

5. (1) 因为 $\forall$ 方阵 $\boldsymbol{A}$，有 $\boldsymbol{A}=\boldsymbol{I}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{I}$，所以 $\boldsymbol{A}\simeq\boldsymbol{A}$.

(2) 若 $\boldsymbol{A}\simeq\boldsymbol{B}$，则 $\exists$ 可逆矩阵 $\boldsymbol{C}$，使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B}$，则 $\boldsymbol{A}=(\boldsymbol{C}^{\mathrm{T}})^{-1}\boldsymbol{B}\boldsymbol{C}^{-1}$，即 $\boldsymbol{B}\simeq\boldsymbol{A}$.

(3) 若 $\boldsymbol{A}\simeq\boldsymbol{B}$，则 $\exists$ 可逆矩阵 $\boldsymbol{C}_1$，使得 $\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}_1=\boldsymbol{B}$.

$\boldsymbol{B}\simeq\boldsymbol{C}$，则 $\exists$ 可逆矩阵 $\boldsymbol{C}_2$，使得 $\boldsymbol{C}_2^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}_2=\boldsymbol{C}$.

所以 $\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{C}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}_2\boldsymbol{C}_1=(\boldsymbol{C}_2\boldsymbol{C}_1)^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{C}_2\boldsymbol{C}_1)=\boldsymbol{C}$.从而 $\boldsymbol{A}\simeq\boldsymbol{C}$.

6. 设矩阵 $\boldsymbol{A}$ 的秩为 r，则 $\exists$ 可逆矩阵 $\boldsymbol{C}$，使得 $\boldsymbol{A}=\boldsymbol{C}^{\mathrm{T}}\left(\begin{array}{c:c}\boldsymbol{D}_r & \boldsymbol{O}\\ \hdashline \boldsymbol{O} & \boldsymbol{O}\end{array}\right)\boldsymbol{C}$.

其中，$\boldsymbol{D}_r=\begin{pmatrix}d_1 & & \boldsymbol{O}\\ & \ddots & \\ \boldsymbol{O} & & d_r\end{pmatrix}$为秩为 r 的对角矩阵.

$$所以\ \boldsymbol{A}=\boldsymbol{C}^{\mathrm{T}}\begin{pmatrix}d_1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0\end{pmatrix}\boldsymbol{C}+\boldsymbol{C}^{\mathrm{T}}\begin{pmatrix}0 & & & \\ & d_2 & & \\ & & 0 & \\ & & & \ddots & \\ & & & & 0\end{pmatrix}\boldsymbol{C}+\cdots$$

$$+C^{\mathrm{T}}\begin{pmatrix}0&&&&&&\\&\ddots&&&&&\\&&0&&&&\\&&&d_r&&&\\&&&&0&&\\&&&&&\ddots&\\&&&&&&0\end{pmatrix}C.$$

其中,每一个加项均为秩是 1 的对称矩阵.

从而原命题成立.

7. 因为 k 是实对称矩阵 $\boldsymbol{A}$ 对应特征值,则 $\exists\, k$ 的特征向量 $\boldsymbol{\alpha}\in\mathbb{R}^n$,且 $\boldsymbol{\alpha}\neq\boldsymbol{0}$,$\boldsymbol{A\alpha}=k\boldsymbol{\alpha}$,两边同时左乘 $\boldsymbol{\alpha}^{\mathrm{T}}$,$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A\alpha}=\boldsymbol{\alpha}^{\mathrm{T}}k\boldsymbol{\alpha}=k\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}$.

8. (1) $\boldsymbol{A}=\begin{pmatrix}-5&0&0\\0&3&0\\0&0&-2\end{pmatrix}$,所以 $r(\boldsymbol{A})=3$.从而它在复数域上的规范形为 $y_1^2+y_2^2+y_3^2$.

所作非退化线性替换为 $\boldsymbol{X}=\boldsymbol{CY}$,其中 $\boldsymbol{C}=\begin{pmatrix}\frac{1}{\sqrt{-5}}&&\\&\frac{1}{\sqrt{3}}&\\&&\frac{1}{\sqrt{-2}}\end{pmatrix}=\begin{pmatrix}-\frac{\sqrt{5}}{5}\mathrm{i}&&\\&\frac{\sqrt{3}}{3}&\\&&\frac{\sqrt{2}}{2}\mathrm{i}\end{pmatrix}$.

(2) $\boldsymbol{A}=\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\end{pmatrix}$,下面求一矩阵 $\boldsymbol{C}_1$,使得 $\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{AC}_1$ 为对角矩阵.

$$\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}2&1&-2\\1&0&-3\\-2&-3&0\\1&0&0\\1&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}2&0&0\\1&-\frac{1}{2}&-2\\-2&-2&-2\\1&-\frac{1}{2}&1\\1&\frac{1}{2}&1\\0&0&1\end{pmatrix}$$

$$\to\begin{pmatrix}2&0&0\\0&-\frac{1}{2}&-2\\0&-2&-2\\1&-\frac{1}{2}&1\\1&\frac{1}{2}&1\\0&0&1\end{pmatrix}\to\begin{pmatrix}2&0&0\\0&-\frac{1}{2}&0\\0&0&6\\1&-\frac{1}{2}&3\\1&\frac{1}{2}&-1\\0&0&1\end{pmatrix}.$$

所以 $\boldsymbol{C}_1=\begin{pmatrix}1&-\frac{1}{2}&3\\1&\frac{1}{2}&-1\\0&0&1\end{pmatrix}$，且 $\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix}2&&\\&-\frac{1}{2}&\\&&6\end{pmatrix}$.

令 $\boldsymbol{C}=\boldsymbol{C}_1\begin{pmatrix}\frac{1}{\sqrt{2}}&&\\&\frac{1}{\sqrt{-\frac{1}{2}}}&\\&&\frac{1}{\sqrt{6}}\end{pmatrix}=\begin{pmatrix}\frac{\sqrt{2}}{2}&\frac{\sqrt{2}}{2}\mathrm{i}&\frac{\sqrt{6}}{2}\\\frac{\sqrt{2}}{2}&-\frac{\sqrt{2}}{2}\mathrm{i}&-\frac{\sqrt{6}}{6}\\0&0&\frac{\sqrt{6}}{6}\end{pmatrix}$，

则 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{I}$.

从而 $f(x_1,x_2,x_3)$ 在复数域上的规范形为 $y_1^2+y_2^2+y_3^2$. 所作非退化线性替换为 $\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$.

(3) $\boldsymbol{A}=\begin{pmatrix}0&\frac{-1-\mathrm{i}}{2}\\\frac{-1-\mathrm{i}}{2}&2\mathrm{i}\end{pmatrix}$ 下面求一矩阵 $\boldsymbol{C}_1$，使得 $\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}_1$ 为对角矩阵.

$$\begin{pmatrix}0&\frac{-1-\mathrm{i}}{2}\\\frac{-1-\mathrm{i}}{2}&2\mathrm{i}\\1&0\\0&1\end{pmatrix}\to\begin{pmatrix}-1-\mathrm{i}&\frac{-1+3\mathrm{i}}{2}\\\frac{-1+3\mathrm{i}}{2}&2\mathrm{i}\\1&0\\1&1\end{pmatrix}\to\begin{pmatrix}-1-\mathrm{i}&0\\0&\frac{-7+5\mathrm{i}}{2}\\1&\frac{1+2\mathrm{i}}{2}\\1&\frac{3+2\mathrm{i}}{2}\end{pmatrix},$$

所以 $C_1=\begin{pmatrix}1 & \dfrac{1+2i}{2}\\ 1 & \dfrac{3+2i}{2}\end{pmatrix}$，且 $C_1^{T}AC_1=\begin{pmatrix}-1-i & 0\\ 0 & \dfrac{-7+5i}{2}\end{pmatrix}$.

$$令\ C=C_1\begin{pmatrix}\dfrac{1}{\sqrt{-1-i}} & 0\\ 0 & \dfrac{\sqrt{2}}{\sqrt{-7+5i}}\end{pmatrix}=\begin{pmatrix}\dfrac{1}{\sqrt{-1-i}} & \dfrac{1+2i}{2}\sqrt{\dfrac{2}{-7+5i}}\\ \dfrac{1}{\sqrt{-1-i}} & \dfrac{3+2i}{2}\sqrt{\dfrac{2}{-7+5i}}\end{pmatrix},$$

则 $C^{T}AC=I$.

从而 $f(x_1,x_2)$ 在复数域上的规范形为 $y_1^2+y_2^2$.所作的非退化线性替换为 $X=CY$.

9. (1) $f(x_1,x_2,x_3)$ 在实数域上的规范形为 $-y_1^2+y_2^2-y_3^2$.所作的非退化线性替换为 $X=CY$,

其中 $C=\begin{pmatrix}\dfrac{\sqrt{5}}{5} & & \\ & \dfrac{\sqrt{3}}{3} & \\ & & \dfrac{\sqrt{2}}{2}\end{pmatrix}$.秩为 3,正惯性指数为 1,负惯性指数为 2,符号差为−1.

(2) $f(x_1,x_2,x_3)$ 在实数域上的规范形为 $y_1^2-y_2^2+y_3^2$.所作的非退化线性替换为 $X=CY$.

其中,$C=\begin{pmatrix}1 & -\dfrac{1}{2} & 3\\ 1 & \dfrac{1}{2} & -1\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\dfrac{1}{\sqrt{2}} & & \\ & \sqrt{2} & \\ & & \dfrac{1}{\sqrt{6}}\end{pmatrix}=\begin{pmatrix}\dfrac{\sqrt{2}}{2} & -\dfrac{\sqrt{2}}{2} & \dfrac{\sqrt{6}}{2}\\ \dfrac{\sqrt{2}}{2} & \dfrac{\sqrt{2}}{2} & -\dfrac{\sqrt{6}}{6}\\ 0 & 0 & \dfrac{\sqrt{6}}{6}\end{pmatrix}$.秩为 3,正惯性指数为 1,

负惯性指数为 2,符号差为−1.

(3) $f(x_1,x_2,x_3,x_4)$ 在实数域上的规范形为 $-y_1^2+y_2^2+y_3^2$.所作非退化线性替换为

$Y=XC$,其中 $C=\begin{pmatrix}1 & 2 & -2\\ -2 & 0 & 4\\ 3 & -1 & -6\\ -1 & 3 & 2\end{pmatrix}$秩为 3,正惯性指数为 2,负惯性指数为 1,符号差为 1.

10. 可分为 10 类.

① 秩为 1:$\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$,$\begin{pmatrix}-1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$.

② 秩为 2:$\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}$,$\begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix}$,$\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&0\end{pmatrix}$.

③ 秩为 3:$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$,$\begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix}$,$\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix}$,$\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&-1\end{pmatrix}$.

④ 秩为 0:$\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}$.

一般地,n 阶实对称矩阵组成的集合按照合同关系,可以分为 $1+2+3+4+\cdots+(n+1)$ $=\dfrac{(n+1)(n+2)}{2}$个类.

11. 因为 $\exists\,\boldsymbol{\alpha}_1\in\mathbb{R}^n$,使得 $f(\boldsymbol{\alpha}_1)=\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}_1>0$,所以 f 不是半负定的.

又因为 $\exists\,\boldsymbol{\alpha}_2\in\mathbb{R}^n$,使得 $f(\boldsymbol{\alpha}_2)=\boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}_2<0$,所以 f 不是半正定的,从而 f 是不定的.

从而 f 可以通过非退化线性替换 $\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$ 化为规范形:

$f(\boldsymbol{X})=y_1^2+y_2^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2\ (1\leqslant p<r)$.

不妨取 $\boldsymbol{\alpha}_4=(1,0,\cdots,0,1)$,$\boldsymbol{\alpha}_3=\boldsymbol{C}\boldsymbol{\alpha}_4\neq\mathbf{0}$,则有 $f(\boldsymbol{\alpha}_3)=\boldsymbol{\alpha}_3^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}_3=0$.

12. 设 $f(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 为正定二次型,$g(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{X}$ 为与 $f(\boldsymbol{X})$ 等价的二次型,则正定矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 合同.故 $\exists$ 可逆矩阵 $\boldsymbol{Q}$,使得 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{B}$.

又因为 $\boldsymbol{A}$ 正定,所以 $\exists$ 可逆矩阵 $\boldsymbol{P}$,使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}=\boldsymbol{A}$.

所以 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{Q}=\boldsymbol{B}$,即 $(\boldsymbol{P}\boldsymbol{Q})^{\mathrm{T}}(\boldsymbol{P}\boldsymbol{Q})=\boldsymbol{B}$.令 $\boldsymbol{C}=\boldsymbol{P}\boldsymbol{Q}$,则 $\boldsymbol{C}$ 可逆.

且 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}=\boldsymbol{B}$,所以 $\boldsymbol{B}$ 正定,进而 $g(\boldsymbol{X})$ 也正定.

13. 设 $\boldsymbol{A}_n,\boldsymbol{B}_n\in M_n(\mathbb{R})$,且 $\boldsymbol{A},\boldsymbol{B}$ 半正定.

故 $\forall\,\boldsymbol{\alpha}\in\mathbb{R}^n(\boldsymbol{\alpha}\neq 0)$,均有 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}\geqslant 0$,$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\alpha}\geqslant 0$.

所以 $\boldsymbol{\alpha}^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{\alpha}=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}+\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\alpha}\geqslant 0$,故 $\boldsymbol{A}+\boldsymbol{B}$ 半正定.

若其中之一为正定.不妨设 $\boldsymbol{B}$ 正定,则 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\alpha}>0$.

所以 $\boldsymbol{\alpha}^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{\alpha}=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}+\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\alpha}>0$,故 $\boldsymbol{A}+\boldsymbol{B}$ 正定.

14. 设 $\boldsymbol{A}_n\in M_n(\mathbb{R})$,且 $\boldsymbol{A}$ 正定,则 $\exists$ 可逆矩阵 $\boldsymbol{C}$,使得

$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{I}$.所以 $\boldsymbol{A}=(\boldsymbol{C}^{\mathrm{T}})^{-1}\boldsymbol{I}\boldsymbol{C}^{-1}$.

进而 $\boldsymbol{A}^{-1}=\boldsymbol{C}\boldsymbol{I}\boldsymbol{C}^{\mathrm{T}}$,从而 $\boldsymbol{A}^{-1}\simeq\boldsymbol{I}$.故 $\boldsymbol{A}^{-1}$ 正定.

15. 若 $\boldsymbol{A}$ 半正定,则 $\forall\,\boldsymbol{\alpha}\in F^n$,$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\alpha}\geqslant 0$.

充分性:假设它的正惯性指数 $p\neq$ 秩 r 则 $p<r$.则规范形为 $y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2$.

设所作非退化替换为 $\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y}$.

取 $\boldsymbol{\alpha}=(0,0,\cdots,0,1)$,则 $(\boldsymbol{C}^{-1}\boldsymbol{\alpha})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{C}^{-1}\boldsymbol{\alpha})=-1<0$,矛盾.

故 $p=r$.

必要性显然.故充要条件成立.

同理可证负定、半负定的充要条件.

16. (1) 该二次型的矩阵为 $A=\begin{pmatrix}99&-6&24\\-6&130&-30\\24&-30&71\end{pmatrix}$.

顺序主子式 $|A_1|>0$, $|A_2|=\begin{vmatrix}99&-6\\-6&130\end{vmatrix}>0$, $|A_3|>0$.

所以该二次型是正定的.

(2) 该二次型的矩阵为:$A=\begin{pmatrix}10&4&12\\4&2&-6\\12&-6&3\end{pmatrix}$.

顺序主子式 $|A_1|=1>0$, $|A_2|=4>0$, $|A_3|=-1\ 212<0$.

所以该二次型不是正定的.

17. (1) 二次型 $f(x_1,x_2,x_3)=x_1^2+x_2^2+5x_3^2+2tx_1x_2-2x_1x_3+4x_2x_3$ 的矩阵为

$$A=\begin{pmatrix}1&t&-1\\t&1&2\\-1&2&5\end{pmatrix}.$$

所以 $|A_1|=1>0$, $|A_2|=\begin{vmatrix}1&t\\t&1\end{vmatrix}=1-t^2$,

$$|A_3|=\begin{vmatrix}1&t&-1\\t&1&2\\-1&2&5\end{vmatrix}=\begin{vmatrix}1&t&-1\\t+2&1+2t&0\\4&2+5t&0\end{vmatrix}=-\begin{vmatrix}t+2&1+2t\\4&2+5t\end{vmatrix}=-(5t^2+4t).$$

要使 $f(x_1,x_2,x_3)$ 正定,则 $\begin{cases}|A_2|>0,\\|A_3|>0,\end{cases}$ 即 $\begin{cases}1-t^2>0,\\-(5t^2+4t)>0\end{cases}\Rightarrow -\dfrac{4}{5}<t<0$.

(2) 二次型 $f(x_1,x_2,x_3)=x_1^2+4x_2^2+x_3^2+2tx_1x_2+2x_2x_3$ 的矩阵为

$$A=\begin{pmatrix}1&t&0\\t&4&1\\0&1&1\end{pmatrix}.$$

所以 $|A_1|=1>0$, $|A_2|=\begin{vmatrix}1&t\\t&4\end{vmatrix}=4-t^2$, $|A_3|=\begin{vmatrix}1&t&0\\t&4&1\\0&1&1\end{vmatrix}=\begin{vmatrix}4&1\\1&1\end{vmatrix}-t\begin{vmatrix}t&1\\0&1\end{vmatrix}=3-t^2$.

要使 $f(x_1,x_2,x_3)$ 正定,则 $\begin{cases}|A_2|>0,\\|A_3|>0,\end{cases}$ 即 $\begin{cases}4-t^2>0,\\3-t^2>0\end{cases}\Rightarrow -\sqrt{3}<t<\sqrt{3}$.

18. "$\Rightarrow$":若 A 正定,则∃正交矩阵 Q 和对角元全为正数的对角矩阵 D,使得 $A=Q^{\mathrm{T}}DQ$.

记 P 是对角阵,且其对角元是 D 的对角元的平方根,即 P 满足 $P^2=D$,则 $A=$

$Q^{T}P^{T}PQ=(PQ)^{T}PQ$.记 $C=PQ$,则 $A=C^{T}C=C^{2}$.

"$\Leftarrow$":若 $A=C^{T}C=C^{2}$,则 $\forall\, \mathbf{0}\neq\boldsymbol{\alpha}\in F^{n}$,有

$\boldsymbol{\alpha}^{T}A\boldsymbol{\alpha}=\boldsymbol{\alpha}^{T}C^{T}C\boldsymbol{\alpha}=(C\boldsymbol{\alpha})^{T}(C\boldsymbol{\alpha})=|C\boldsymbol{\alpha}|^{2}>0$.

所以 A 正定.

19. 因为 A 正定$\Leftrightarrow$$A$ 的所有顺序主子式都大于零,而且任一 A 的合同矩阵也正定.

又因为 A 的任一主子式 K 均可以找到一个与 A 合同的矩阵 B,使得 K 为 B 的顺序主子式(由初等行列变换法易知).

从而 A 的所有主子式都大于零.

若 A 为实对称正定矩阵,则它的主对角线上元素均为一阶主子式,故大于 0,为正数.

20. (1) 对矩阵 A 的阶数进行归纳.

当 $n=1$ 时,结论自然成立.

假设当 $n=k-1(k\in\mathbb{N}_{+})$时结论成立,即 $A\in M_{n-1}(\mathbb{R})$正定,则 $|A|\leqslant a_{11}a_{22}\cdots a_{n-1,n-1}$,

那么,当 $n=k(k\in\mathbb{N}_{+})$时,记 $A=\begin{pmatrix}P_{n-1} & B\\ B^{T} & a_{nn}\end{pmatrix}$,其中 $P_{n-1}=\begin{pmatrix}a_{11} & \cdots & a_{1,n-1}\\ \vdots & & \vdots\\ a_{n-1,1} & \cdots & a_{n-1,n-1}\end{pmatrix}$,

$B=\begin{pmatrix}a_{1n}\\ \vdots\\ a_{n-1,n}\end{pmatrix}$.

令 $F=\begin{pmatrix}I_{n-1} & -P_{n-1}^{-1}B\\ O & 1\end{pmatrix}$,则 $F^{T}AF=\begin{pmatrix}P_{n-1} & O\\ O & a_{nn}-B^{T}P_{n-1}^{-1}B\end{pmatrix}$.

由于 P_{n-1} 正定,所以 $\exists$ 可逆矩阵 C,使得 $C^{T}P_{n-1}C=I$.从而 $C^{-1}P_{n-1}^{-1}(C^{T})^{-1}=C^{-1}P_{n-1}^{-1}(C^{-1})^{T}=I$.

因而 P_{n-1}^{-1}是正定的.从而 $B^{T}P_{n-1}^{-1}B\geqslant 0$.

又因为 $|F|=|F^{T}|=1$,所以 $|A|=|F^{T}||A||F|=|F^{T}AF|=|P_{n-1}||a_{nn}-B^{T}P_{n-1}^{-1}B|$.

从而 $|A|>0$,$|P_{n-1}|>0$,从而 $a_{nn}-B^{T}P_{n-1}^{-1}B>0$.

即 $0\leqslant B^{T}P_{n-1}^{-1}B\leqslant a_{nn}$.

所以 $|A|\leqslant|P_{n-1}|\cdot a_{nn}\leqslant a_{11}a_{12}\cdots a_{nn}$.

故 $\forall\, n\in\mathbb{N}_{+}$,$|A|\leqslant a_{11}a_{22}\cdots a_{nn}$成立.

(2) 若 $|T|=0$. 则不等式成立.

若 $|T|\neq 0$. 则 $T^{T}T$ 是对称矩阵且正定.

而 $T^{T}T$ 的对角线上的元素 $a_{ii}=t_{1i}^{2}+t_{2i}^{2}+\cdots+t_{ni}^{2}$.

由(1)可得

$$|T^{T}T|\leqslant\prod_{i=1}^{n}a_{ii}=\prod_{i=1}^{n}(t_{1i}^{2}+t_{2i}^{2}+\cdots+t_{ni}^{2}),$$

所以 $|\boldsymbol{T}|=\sqrt{|\boldsymbol{T}^{\mathrm{T}}\boldsymbol{T}|}\leqslant\prod_{i=1}^{n}\sqrt{t_{1i}^2+t_{2i}^2+\cdots+t_{ni}^2}$.

即原命题成立.

21. 若 $\boldsymbol{A}$ 负定,则 $-\boldsymbol{A}$ 正定.进而 $-\boldsymbol{A}$ 的奇数阶顺序主子式 $|-\boldsymbol{A}_k|$ 均为正.

所以 $|\boldsymbol{A}_k|<0$,即 $\boldsymbol{A}$ 的奇数阶顺序主子式全为负数.即 $|-\boldsymbol{A}_k|=(-1)^k|\boldsymbol{A}_k|>0$.

同理,$-\boldsymbol{A}$ 的偶数阶顺序主子式为正,即 $|-\boldsymbol{A}_m|=(-1)^m|\boldsymbol{A}_m|>0$($m$ 为偶数).

从而 $|\boldsymbol{A}_m|>0$,即 $\boldsymbol{A}$ 的偶数阶顺序主子式全为正数.

习 题 八

1. 因为 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}$ 为一个数,所以 $(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y})\boldsymbol{y}=(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y})\boldsymbol{I}\boldsymbol{y}=\boldsymbol{y}(\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x})\boldsymbol{I}=(\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}})\boldsymbol{x}$.

2. (1) $\dfrac{1}{25}\begin{pmatrix}7&24\\24&-7\end{pmatrix}$;(2) $\dfrac{1}{25}\begin{pmatrix}16&-12\\-12&9\end{pmatrix}$;(3) $\dfrac{1}{2}\begin{pmatrix}1&-\sqrt{3}\\\sqrt{3}&1\end{pmatrix}$;(4) $\dfrac{1}{9}\begin{pmatrix}8&-2&2\\-2&5&4\\2&4&5\end{pmatrix}$.

3. 对 $\boldsymbol{x}\in\mathbb{R}^n$,设 $\boldsymbol{x}$ 可以分解为 $\boldsymbol{x}=a_1\boldsymbol{\alpha}_1+\cdots+a_r\boldsymbol{\alpha}_r+\boldsymbol{\beta}$,其中 $\boldsymbol{\beta}$ 与 $\boldsymbol{P}$ 正交,则不难验证 $a_i=(\boldsymbol{x},\boldsymbol{\alpha}_i)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\alpha}_i$.设 $\boldsymbol{x}$ 对 $\boldsymbol{P}$ 的正交投影变换为 φ,则

$\varphi(\boldsymbol{x})=a_1\boldsymbol{\alpha}_1+\cdots+a_r\boldsymbol{\alpha}_r=(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\alpha}_1)\boldsymbol{\alpha}_1+\cdots+(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\alpha}_r)\boldsymbol{\alpha}_r=(\boldsymbol{\alpha}_1\boldsymbol{\alpha}_1^{\mathrm{T}})\boldsymbol{x}+\cdots+(\boldsymbol{\alpha}_r\boldsymbol{\alpha}_r^{\mathrm{T}})\boldsymbol{x}$.

令 $\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)$ 为 $n\times r$ 矩阵,则 $\varphi(\boldsymbol{x})=\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{x}$,即 φ 对应的矩阵为 $\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}$.从而 $\boldsymbol{x}$ 关于 V 的对称向量为 $2\varphi(\boldsymbol{x})-\boldsymbol{x}$,对应的矩阵为 $2\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}-\boldsymbol{I}$.

2 维、3 维空间中公式的重新推导略.

4. $\begin{pmatrix}40&-27\\55&-37\end{pmatrix}$.

5. (1) 1,2,1,2;(2) 2,1;(3) $\begin{pmatrix}1\,026\\2\,050\end{pmatrix}$.

6. 判断正误:

(1) 错误,考虑伸缩变换.

(2) 错误,同上.

(3) 正确.

(4) 正确.

(5) 正确.

(6) 正确.

(7) 正确.

(8) 正确.

(9) 正确.

(10) 正确.

(11) 错误,$\mathbb{R}^2$中的旋转.

(12) 错误,$\mathbb{R}^2$中的旋转.

(13) 错误,$\mathbb{R}^2$中关于原点对称.

(14) 正确,事实上该结论对一般的 n 维欧氏空间都成立.

(15) 错误,$\mathbb{R}^2$中关于 x 轴对称.

7. 略.

8. 略.

9. 因为 $\boldsymbol{v}\perp W$,又 $\boldsymbol{v}\in W$,所以$(\boldsymbol{v},\boldsymbol{v})=0$,由于内积的正定性,有 $\boldsymbol{v}=\boldsymbol{0}$.

10. 有解,设 $\boldsymbol{A}=(a,b,c)$,$d=-1$,则该方程等价于 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=\boldsymbol{A}^{\mathrm{T}}d$.

11. 图略,回归矩阵方程 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}$ 为 $\begin{pmatrix}1&3\\1&4\\1&5\\1&6\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}=\begin{pmatrix}2.5\\3\\4\\4.5\end{pmatrix}$,且 $\boldsymbol{A}$ 列满秩.则 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=$

$\begin{pmatrix}1&1&1&1\\3&4&5&6\end{pmatrix}\begin{pmatrix}1&3\\1&4\\1&5\\1&6\end{pmatrix}=\begin{pmatrix}4&18\\18&86\end{pmatrix}$, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}=\begin{pmatrix}14\\66.5\end{pmatrix}$, $\begin{pmatrix}\hat{a}\\\hat{b}\end{pmatrix}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})=$

$\dfrac{1}{20}\begin{pmatrix}86&-18\\-18&4\end{pmatrix}\begin{pmatrix}14\\66.5\end{pmatrix}=\begin{pmatrix}0.35\\0.7\end{pmatrix}$. 所以回归方程为 $\hat{y}=0.35+0.7x$.

12. 185 cm.

*13. (1) 由题意 $\boldsymbol{A}=\begin{pmatrix}1&x_1\\1&x_2\\\vdots&\vdots\\1&x_n\end{pmatrix}$,$\boldsymbol{y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$,那么

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix}n&\sum x_i\\\sum x_i&\sum x_i^2\end{pmatrix},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}=\begin{pmatrix}\sum y_i\\\sum x_iy_i\end{pmatrix},$$

$$D=\det\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=n\sum x_i^2-\left(\sum x_i\right)^2.$$

故而

$$\begin{pmatrix}a\\b\end{pmatrix}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y})=\frac{1}{D}\begin{pmatrix}\sum x_i^2&-\sum x_i\\-\sum x_i&n\end{pmatrix}\begin{pmatrix}\sum y_i\\\sum x_iy_i\end{pmatrix}$$

$$=\frac{1}{D}\begin{pmatrix}\sum x_i^2\sum y_i-\sum x_i\sum x_iy_i\\n\sum x_iy_i-\sum x_i\sum y_i\end{pmatrix}.$$

(2) 直接在最后一步以 x_i+c 代替 x_i. 则有

$$\begin{aligned} D' &= n\sum(x_i+c)^2-(\sum(x_i+c))^2 \\ &= n\sum x_i^2+2nc\sum x_i+n^2c^2-(\sum x_i+nc)^2 \\ &= n\sum x_i^2-(\sum x_i)^2=D\ , \end{aligned}$$

且有

$$\begin{aligned} D(b'-b) &= (n\sum(x_i+c)y_i-\sum(x_i+c)\sum y_i)-(n\sum x_iy_i-\sum x_i\sum y_i) \\ &= n\sum cy_i-\sum c\sum y_i=nc\sum y_i-nc\sum y_i=0, \end{aligned}$$

故有 $b'=b$.

$$\begin{aligned} D(a'-a) &= (\sum(x_i+c)^2\sum y_i-\sum(x_i+c)\sum(x_i+c)y_i) \\ &\quad -(\sum x_i^2\sum y_i-\sum x_i\sum x_iy_i) \\ &= \sum\sum(2cx_iy_j+c^2y_j)-\sum\sum(c(x_i+x_j+c)y_j) \\ &= \sum\sum(cx_iy_j-cx_jy_j)=c\sum x_i\sum y_i-nc\sum x_iy_i=-cDb, \end{aligned}$$

故有 $a'=a-bc$. 证毕.

14. 系数矩阵为 $\widetilde{A}=\begin{pmatrix}6&-2&-2\\-2&3&-4\\-2&-4&3\end{pmatrix}$,算出 $A=\begin{pmatrix}6&-2\\-2&3\end{pmatrix}$特征值为 $\lambda_1=7,\lambda_2=2$,算出其特征向量,并单位正交化,得 $C=\begin{pmatrix}\frac{2}{\sqrt{5}}&\frac{1}{\sqrt{5}}\\\frac{-1}{\sqrt{5}}&\frac{2}{\sqrt{5}}\end{pmatrix}$. 令 $X=CY$,得 $7y_1^2+2y_2^2-4\sqrt{5}\,y_2+3=0$. 令 $\begin{pmatrix}z_1\\z_2\end{pmatrix}=\begin{pmatrix}y_1\\y_2-\sqrt{5}\end{pmatrix}$, 即 $Y=Z+\begin{pmatrix}0\\\sqrt{5}\end{pmatrix}$,得 $z_1^2+\frac{z_2^2}{7/2}=1$.

15. 系数矩阵为 $A=\begin{pmatrix}15&3&-1\\3&11&-2\\-1&-2&9\end{pmatrix}$,显然 $I_1,I_2>0$,这是抛物线型曲线. 经验证,有 $I_3=\det A>0,I_3/I_1>0$(也可根据 A 是一个严格对角占优阵得上述结论).故而这是虚椭圆.

致谢

编者首先特别感谢北京师范大学张英伯教授,她对本书的课程定位、选材、编排次序以及细节校订等诸多方面提出了很多指导性的意见和建议,使本书增色不少。

编者衷心感谢中国大学先修课程管理委员会课程研发与教学管理办公室的侯进科老师、车晓佳老师和黄平老师,是他们的精心组织、周密协调,才使本教材能最终呈现给大家。编者还要感谢清华大学附属中学的范永春老师、中国人民大学附属中学的李海波老师、北京市十一学校屈楠老师、辽宁实验中学的王清礼老师等,他们也对本教材提出过很好的建议,并无私地贡献了他们的教学体会和感受。

编者感谢高等教育出版社数学分社的支持,特别要感谢李茜编辑的努力。她多次参加了我们的研讨,提出了许多富有建设性的建议并组织了细致认真的编辑加工,使本书可以展现在读者面前。

编者还要感谢附表中所有老师!他们参加了前几期的中国大学先修课程试点项目实验学校教师培训以及 MOOCAP 教师工作坊,各位老师认真研读了讲义的相关章节,并提出了很多中肯的意见和建议。他们的许多建议已经体现在本书目前的版本中,我们希望能够继续得到他们以及更多中学一线教师的关注,欢迎大家提出更多更新的修改意见及建议。

编　者

2016 年 11 月

附表 1　2014 年 8 月中国大学先修课程试点项目首批实验学校教师培训参训名单

陈银科	合肥市第一中学	邹素文	柳州高级中学
屈楠	北京市十一学校	唐毓仁	广西南宁市第二中学
刘舒眉	清华大学附属中学	禹邵宁	广西南宁市第二中学
李海波	中国人民大学附属中学	景燕波	石家庄二中
蒋燕敏	福州第一中学	程建辉	河南省实验中学
张婷	西北师范大学附属中学	张玉	郑州外国语学校
魏自盈	西北师范大学附属中学	王久成	东北师范大学附属中学

续表

徐盛馀	江苏省天一中学	李青	成都七中嘉祥外国语学校
王清礼	辽宁省实验中学	黄波	成都市树德中学
汪建国	银川唐徕回民中学	张世军	成都市树德中学
杨伟	山西大学附属中学校	舒召龙	四川省南充高级中学
石鹏	西安交通大学附属中学	潘铁	天津市实验中学
赵钊	西安市铁一中学	卢翔	天津市耀华中学
刘雷	西北工业大学附属中学	南天时	天津市耀华中学
侯磊	上海交通大学附属中学	高荣	新疆兵团第二师华山中学

附表 2　2015 年 1 月中国大学先修课程试点项目实验学校第二期教师培训参训名单

何震	东北师范大学附属中学	虞哲骏	宁波市镇海中学
行凯歌	江苏省天一中学	张义斌	浙江省宁波市镇海中学
曾文军	重庆市第八中学校	党效文	西北工业大学附属中学
袁莹	河南省实验中学	潘克亮	浙江省温州中学
李喆	上海交通大学附属中学	郭静静	北京市第一〇一中学
李维佳	西安交通大学附属中学	张文	大连市第二十四中学
管培祥	珠海市第一中学	程竹	北京景山学校
宛昭勋	石家庄市第二中学	钱铭	无锡市第一中学
王晓明	昌邑市第一中学	李凤	北京师范大学附属中学
王亚力	山西大学附属中学校	张玉	郑州外国语学校
许桂兵	成都外国语学校	徐笑	上海市格致中学
张婷	西北师范大学附属中学	孙理	中国人民大学附属中学
林海	福建省晋江市第一中学	徐丹	河北衡水中学

附表 3　2015 年 7 月中国大学先修课程试点项目实验学校第三期教师培训参训名单

黄建	电子科技大学实验中学	高磊	山东省青岛第二中学
何震	东北师范大学附属中学	吴文庆	中国人民大学附属中学
林芬芳	福建省晋江市第一中学	刘刚	山东省昌邑市文山中学
周培祥	安徽合肥一六八中学	孙桂云	山东省昌邑市文山中学
肖奎	河南省实验中学	杨伟	山西大学附属中学校
姜宗帅	河北衡水中学	屈楠	北京市十一学校
孙承辉	江苏省天一中学	宛昭勋	河北省石家庄二中
安恺凯	江苏省天一中学	任大伟	河北省石家庄一中
王清礼	辽宁省实验中学	杨晓文	四川省什邡市七一中学
王东刚	山东青岛市第五十八中学	钱铭	江苏省无锡市第一中学

续表

李维佳	西安交通大学附属中学	刘东	西北工业大学附属中学
高荣	新疆兵团第二师华山中学	高明月	西安交通大学苏州附属中学
张玉	河南郑州外国语学校	王洁凌	广州市天河外国语学校
张慧敏	湖北省沙市中学	吴艳妹	北京市第十二中学
高维宗	西北师范大学附属中学		

附表 4　中国大学先修课程试点项目实验学校第四期教师培训参训名单

崔文鑫	上海交通大学附属中学	周志杰	北京市第十二中学
张小明	海宁市高级中学	王清礼	辽宁省实验中学
杨晓文	四川省什邡市七一中学	乔晓梅	山东省莱阳市第一中学
张玉	郑州外国语学校	方勇	中山市第一中学
施利强	浙江省德清县高级中学	石因	西安交通大学苏州附属中学
周兰英	福建省泉州第五中学	李云杰	福清市第三中学
邹素文	柳州高级中学	郭铭纪	福建省泉州第一中学
侯界飞	山东省青岛第二中学	祝敏君	福建省闽侯县第一中学
林娟	河北衡水中学	刘永超	吉林省东丰县第二中学
郑林	福州第一中学	屈韬	海南中学
汪林	安徽省铜陵市第一中学	周玉鼎	呼和浩特市第二中学

附表 5　2016 年 3 月 MOOCAP 教师工作坊研讨会参会名单

安博忠	山东广饶一中	荣贺	北京市顺义牛栏山第一中学
曾卫国	长沙市雅礼中学	唐志军	湖南省衡阳市第八中学
陈宏钦	河南省济源一中	王喜朝	南阳市一中
陈庆广	江苏省赣榆高级中学	徐欣欣	浙江省宁波市鄞州中学
陈扬扬	山东省临沂第一中学	闫国栋	郑州市第一中学
高程宇	浙江省乐成寄宿中学	杨伟	山西大学附属中学校
韩亮	营口开发区第一高级中学	姚璐	首都师范大学附属中学
侯伟华	山东省莱芜市第一中学	姚巧莲	东营市胜利第一中学
侯银玲	大庆外国语学校	张利燕	辽宁省盘锦市辽河油田第一高级中学
金春霞	北京市第一七一中学	张淑春	湖北省黄冈中学
孔祥武	常州市第一中学	张旭	山东省高密市第一中学
李海波	中国人民大学附属中学	赵林彪	广西柳州铁一中学
李学涛	河南省安阳市第一中学	赵志远	山东广饶一中
李瑛	浙江省平阳县浙鳌高级中学	郑林	福州一中
陆敬山	上海市崇明中学	姬成双	常州市第一中学
彭玲	东北育才学校	毕嘉宇	清华大学附属中学

附表 6 2016 年 8 月 MOOCAP 教师工作坊研讨会参会名单

赫希武	辽宁省丹东市第二中学	谢海燕	湖南省澧县第一中学
黄建	电子科技大学实验中学	严振宇	吉林省延边第一中学
吕作文	江苏省运河中学	姚才镇	浙江省天台中学
邱运先	广西玉林高级中学	张余婷	上海交通大学附属中学
宋振方	江苏省运河中学	周炜	鄂尔多斯市第一中学
王一	江苏省新海高级中学	周兴奎	鞍山市第一中学
王莹	郑州外国语学校		

郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任；构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人进行严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话　(010) 58581999　58582371　58582488

反盗版举报传真　(010) 82086060

反盗版举报邮箱　dd@hep.com.cn

通信地址　北京市西城区德外大街 4 号
高等教育出版社法律事务与版权管理部

邮政编码　100120

中国大学先修课程

文学写作——格非　陆楠楠　徐慧琳
文学写作作品精读——格非　徐慧琳　陆楠楠
微积分——扈志明　等
概率论与数理统计——叶俊　梁恒　李劲松
线性代数——冯荣权　王殿军　杨晶　周俊
微观经济学——李宏彬　周黎安
物理（力学）——张汉壮
物理（力学）——张留碗　李岩松　俞家新
通用学术英语——吕燕彬

ISBN 978-7-04-048260-7
9 787040 482607 >
定价 48.00元